

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 17:16:04 ; Search time 384.65 Seconds
(without alignments)
16680.383 Million cell updates/sec

Title: US-09-747-371-1

Perfect score: 3737

Sequence: 1 ggcgtccgcacacaccc.....ctaagcactctggagacat 3737

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2204.8	59.0	2386	20	Human breast-speci
2	1944.2	52.0	1949	20	Human breast-speci
3	1900.4	50.9	1922	21	Breast cancer prot
4	1174.4	31.4	1179	21	Breast cancer prot
5	1069.2	28.6	4821	24	Human POLY6 nucleo
6	1067.6	28.6	4550	24	Human POLY7 nucleo
7	959.4	25.7	3177	24	Human POLY8 nucleo
8	940.8	25.2	4440	24	Human POLY5 nucleo
9	921.4	24.7	2982	22	Nucleotide sequenc

10	372.4	24.0	2695	22	AA167200	Nucleotide sequenc
11	363.4	24.0	2897	22	AA167221	Nucleotide sequenc
12	764.8	21.0	811	22	AA144853	Human breast cance
13	629.2	16.8	1464	22	AA178220	Nucleotide sequenc
14	623.4	16.7	2694	22	AA167177	Nucleotide sequenc
15	623.4	16.7	2723	21	AA167045	DNA encoding nove
16	452.4	12.1	463	22	AA117867	Human breast cance
17	428.6	11.5	468	22	AA117782	Human breast cance
18	424.2	11.4	431	22	AA118000	Human breast cance
19	419.6	11.2	448	22	AA110792	Human breast cance
20	412	11.0	576	22	AA110707	Human breast cance
21	408	10.3	416	22	AA117667	Human breast cance
22	383.2	10.3	441	22	AA110592	Human breast cance
23	370.6	9.3	401	22	AA117839	Human breast cance
24	358.2	9.3	1420	22	AA106597	Human cDNA SEQ ID
25	358.2	9.3	1422	22	AA106597	Human cDNA SEQ ID
26	355.8	9.3	2022	22	AA106597	Human cDNA SEQ ID
27	355.8	9.3	2022	22	AA106597	Human cDNA SEQ ID
28	350.6	8.6	415	22	AA110897	Human breast cance
29	320.6	8.6	366	22	AA117442	Human breast cance
30	296	7.9	388	22	AA106550	Human breast cance
31	293.6	7.9	420	22	AA107664	Human breast cance
32	289	7.7	289	20	AA137317	Human breast-speci
33	277	7.4	280	20	AA137312	Human breast-speci
34	276.4	7.4	279	20	AA137316	Human breast-speci
35	276.4	7.4	282	20	AA137314	Human breast-speci
36	273.4	7.3	275	20	AA137306	Human breast-speci
37	270	7.3	270	20	AA137305	Human breast-speci
38	260.4	7.0	1882	21	AA177379	Human cDNA SEQ ID
39	259	6.9	261	20	AA137313	Human breast-speci
40	258.6	6.9	1164	22	AA167199	Nucleotide sequenc
41	257	6.9	291	22	AA17188	Human breast cance
42	257	6.9	291	22	AA171758	Human breast cance
43	254	6.8	254	20	AA137307	Human breast-speci
44	241	6.5	245	20	AA137309	Human breast-speci
45	238	6.1	424	22	AA117908	Human breast cance

ALIGNMENTS

RESULT 1

AA137320

ID: AA137320 standard; DNA: 2386 BP.

XX

AC AA137320:

XX

05-JUL-1999 (first entry)

DI

DE Human breast-specific BS200 DNA EST consensus sequence.

XX

KW Breast: cancer; HS200; EST: expressed sequence tag; human; detection;

KW diagnosis: prevention; treatment: disease predisposition; ss.

XX

OS Homo sapiens.

XX

PN WS902714-AL.

XX

PD 21-JAN-1999.

XX

PF 07-JUL-1998; 95W0-US13908.

XX

PR 07-JUL-1997; 97US-0869127.

XX

PA (ABBO) ABCCT LAB.

XX

PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PH;

PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;

PI Russell JC, Stroupe SD, Yu H;

XX

DR WPI: 1999-120915/10.

XX

PT New breast specific gene HS200 - used to develop products for

[illegible]

192	gaccattatgagagcaagagagcgtgcatttatgccaaatggaaacttcacaaa	251
2072	tgagaaaggacaaatgaacttgtgaaccaatgcccaagaccaggaatattctgggcccgtgaa	2131
252	tgagaaaggacaaatgaacttgtgaaccaatgcccaagaccaggaatattctgggcccgtgaa	311
2132	gaccccaagaacttgaaatgtctgaatgtgaggtctgtgtcaacctgtggaatttc	2191
312	gacccccagagctcgggaactgctcggaatgtcgaggctgtgcaaccactggaatttc	371
2192	tgcaatgggtttgcaactctgcagagctctgtccctgggcacgttccagcctgaagctgg	2251
372	tgcaatgggtttgcaactctgcagagctctgtccctgggcaggttccagcctgaagctgg	431
2252	tcgaacttcctgcttcctctggagagagccttgcaccaaacatcaggagagctacttc	2311
432	tcgaacttcctgcttcctctggagagagccttgcaccaaacatcaggagagctacttc	491
2312	ctttccagagctgtgaaccagagttcaatgttcaactggacattctcaaacacaccac	2371
492	ctttccagagctgtgaaccagagttcaatgttcaactggacattctcaaacacaccac	551
2372	tcaccagatgtattctgtccagatgggaacataccagagctgaatttggaaaaataatg	2431
552	tcaccagatgtattctgtccagatgggaacataccagagctgaatttggaaaaataatg	611
2432	tgtttcttgcagaaataactacagactgaatttgtatgtgtccaaacataaacccagtg	2491
612	tgtttcttgcagaaataactacagactgaatttgtatgtgtccaaacataaacccagtg	671
2492	taaaacagaagaatgtggagggagctggagatctcaactgggtacatttgaatccccaa	2551
672	taaaacagaagaatgtggagggagctggagatctcaactgggtacatttgaatccccaa	731
2552	ctaccaggaactatccagaccacacagagtgtagtggacatataccaccacccccaa	2611
732	ctaccaggaactatccagaccacacagagtgtagtggacatataccaccacccccaa	791
2612	ggcgcgcatctgatcgtgttccttgagatcttctgccatagagagcagcttgggga	2671
792	ggcgcgcatctgatcgtgttccttgagatcttctgccatagagagcagcttgggga	851
2672	ctatcgtgatcggaaaaactctcattccaatctctgtgacaacatatagaaactcgca	2731
852	ctatcgtgatcggaaaaactctcattccaatctctgtgacaacatatagaaactcgca	911
2732	gacctagaacgcccatcgctctcaactccaggtcaaaagaagctgtgagatccagttcaa	2791
912	gacctagaacgcccatcgctctcaactccaggtcaaaagaagctgtgagatccagttcaa	971
2792	gtccaatgaagggaacagcgtatagaggggtccaggtcccatcgtgacatatagaaaga	2851
972	gtccaatgaagggaacagcgtatagaggggtccaggtcccatcgtgacatatagaaaga	1031
2852	ctaccaggaactcattgaagacatagttcgagatggcaggtctatgtatcttggaaacca	2911
1032	ctaccaggaactcattgaagacatagttcgagatggcaggtctatgtatcttggaaacca	1091
2912	tcaggaaatacttaaggaataagaacattcaagagctctatttgatgactcggccatcc	2971
1092	tcaggaaatacttaaggaataagaacattcaagagctctatttgatgactcggccatcc	1151
2972	ccagaactatttcagatcacagcccagggagtcgcgagagatatttccaaagtcttcat	3031
1152	ccagaactatttcagatcacagcccagggagtcgcgagagatatttccaaagtcttcat	1211
3032	ccgattgctacgttccaaaggttccaggtttttgagacttcaaatataactcagcccag	3091
1212	ccgattgctacgttccaaaggttccaggtttttgagacttcaaatataactcagcccag	1271
3092	tgccactcaatacaaatgttctgctataggggttggtgggacagagcttcttcctctgc	3151

[illegible]

us-09-747-371-1.lmg

Wed Jul 3 08:56:42 2002

Db	3254	ATGCAAGATTACAGAGCAGCAGTGAAGTGTAGGGAGTCTTGGGCGTCGAAAGACCA 3195	
Qy	1453	ccctgacgtcgttaagagtggtgagagagaggggtgcttcctcagatgtcctctcgtgca 1512	
Db	3194	TGCTACAGTCAACCGCTGTGGCAGAGAGACACCTGTGCTGCTGACCTGTGCTGCTGAGG 3132	
Qy	1513	ttcaccctcttcagc-----tgtcaccaccctcgtggcaaggttaacatttaagctaa 1566	
Db	3134	CCGATTTTTCAGAGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3075	
Qy	1567	atgaaggaaggttagttgaa-----aaatgctcagctgtttcccgaggtgtcgtgac 1620	
Db	3074	AGGATGCCAATGCCCTTTCGACCTTCCGGAACAAAGGCAAAACAGAGAGAGGCTGGCAGAA 3015	
Qy	1621	cagcactaccagagagacagctcagtaaaagagagcttccgctcagctcagtaaaccttcaat 1680	
Db	3014	CCACAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2955	
Qy	1681	gcagctctggcaagcaagtcctccagagagccctggccagcaagcaccctcaggaagatgt 1740	
Db	2954	GTGACTCTCTCGGAAGGCGCAAGGCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2895	
Qy	1741	ttt---tcactgttgagttgagcttgaactaaacaaagggaggtgacagcttcttctg 1797	
Db	2894	CAAGGCTCACCCTGACTGGAGCGAGAGTTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 2835	
Qy	1798	acctgagctcagctgaagcaagcagagagcgtctcgttaagagccatccgcaagctca 1857	
Db	2834	GGCTGCGCTGCTCGAGAGCGAATGGAAGCGGCGGCTGAAAGAGTCCCTGAAAGTGTCTCA 2775	
Qy	1858	gaagagcctcagagagagagcttcaactcagctctcagctcagctcagctcagctcagct 1917	
Db	2774	GAAAGTCCATCAACAGAGCGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2715	
Qy	1918	ctaaagagcctccagagacatctgaagcgcagga-----gagctcgtggaggtg 1968	
Db	2714	CCACAGCGCGGCTGTGTAGCGGGAGCGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 2655	
Qy	1969	gccaggttcagagaaacaaatgtcaggttcaggttcaggttcaggttcaggttcaggttcag 2028	
Db	2654	GGCAGCAGCGTGTGGGCAACAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2595	
Qy	2029	cacgagaaagctgattttatgtccaaatggaacctttccaaatgaggaagcaaatga 2088	
Db	2594	AGACGGAGCAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2535	
Qy	2089	cttgtaaccatgcccaagacaggaatctctggtggtggtggtggtggtggtggtggtggtg 2148	
Db	2534	CTGGGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2484	
Qy	2149	atatgtctgaatgtgaggtgtgtcgaactgtgtgaattctgtcagatgagcttctgac 2208	
Db	2483	AGTCACACAGTGTGAGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2424	
Qy	2209	cttgcagctctgtgctcgtggcagcttccagctcagctcagctcagctcagctcagctcag 2268	
Db	2423	CTGTGAGCCATGCGCCAGTGGCACTACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2364	
Qy	2269	cctgtgagagagcctgtgcacaaacacacacacacacacacacacacacacacacacacac 2328	
Db	2364	CTTGTGTGGGGCTCTACCAACCAAGATGAAGGGGGGCTCTGTGAGGCGCAGCA 2304	
Qy	2329	ccagatgtaatttcaactgtgacattttacacacacacacacacacacacacacacacacac 2388	
Db	2303	CCAAAGTCCAGTGTCCCGAGGCGCTACTACACACACAGCATCCACCGGTGATTCGT 2244	
Qy	2389	gccagtggaac 2448	
Db	2243	GIGCCATGTGCTCTATGACCCGACCTCCGCTCAGAACTTCGACAGCGGCTGCGAGGAA 2184	
Qy	2449	atactacagactgttattgtctccacaaacacacacacacacacacacacacacacacacac 2508	
Db	2183	ACACAGCAGACAGTGTGTGCTGTACAGTGTGGCCCAATGCAAGAAATCGTCAGTGTG 2124	
Db	2509	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2568	
Db	2123	GIGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 2064	
Qy	2569	caggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2628	
Db	2063	CAGGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 2004	
Qy	2629	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2688	
Db	2003	TGGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1944	
Qy	2689	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2748	
Db	1943	AGGATGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1884	
Qy	2749	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2808	
Db	1883	TGGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1824	
Qy	2809	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2868	
Db	1823	GCGGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1764	
Qy	2869	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2928	
Db	1763	AGGATGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1704	
Qy	2929	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 2988	
Db	1703	ACAAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1644	
Qy	2989	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 3048	
Db	1643	ACAAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1587	
Qy	3049	gaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag 3078	
Db	1586	AGGATGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1557	
RESULT	6		
AB03876			
ID	AB03876	standard: -DNA; 4550 BP.	
XX	AB03876		
AC	AB03876		
DI	14-FEB-2002	(first entry)	
XX	Human POLY A: nucleotide sequence SEQ ID NO:13.		
XX	Human: POLY A: gamma aminobutyric acid receptor; GABA receptor; epidermal growth factor; EGF; complement receptor; HSC; syntaxin; haematopoietic stem and progenitor cell; sulphotransferase; prolactin; antidepressant; neuroleptic; antiparkinsonian; neurotropic; relaxant; anticonvulsant; neuroleptic; neuroprotective; antialcoholic; cardiac; tamoxifen; antihypertensive; psychiatric; medicinal; depression; stroke; Parkinson's disease; Huntington's disease; Tourette's syndrome; anxiety; amyotrophic lateral sclerosis; head trauma; Alzheimer's disease; alcoholism; angina; muscle tension; epileptogenic; memory function; cardiomyopathy; arrhythmogenic right ventricular dysplasia; ss.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FI	CDS	72..2996	
FI		/tag= a	
XX		/product= "POLY7"	
PN	WO200179294-A2		
XX	25-OCT-2001		

Db	1337	tcttgggctctgaaagccatgctcagctgcacccggtctgcaaggaagacacctgtgc	1592
Qy	1493	cctcagatgtcaactctgtggaattcaactctcttcagatgcaca-----ccatcaggac	1516
Db	1397	cctgaactgtccctccaggcccgatcttttgcagaggctcagtgctccattaaaca	1452
Qy	1517	aagtgtaaacttaagctaaatgaagcaagctgtattgttaa-----aaatgcagact	1600
Db	1457	acggcctcctctcaagatcaaggatgcacaaagcgtttgacctgcgaaacaaggcaa	1512
Qy	1601	gtttcccgagggtctgcgaccagcactaccagagaagacacagctcagtataaagagact	1650
Db	1517	aacagaggagctggcagaaacacagggcggtgtgcctctctctgaatgccagct	1576
Qy	1661	cgtctacgtataactacatgcagctctgcgaagcaagtcaccagagccctgtgcgcac	1720
Db	1577	caacttcacacttaagtgtactctctctctgcgaaggcgaagccgcagcggccggac	1636
Qy	1721	aagcacccctaaagaaatttta---tcactgttgagtgtgagcttgaaactaacbaaa	1777
Db	1637	ccctccaggcagaagggtgcacaggctcacctggaactggaagcagaggtctcagcga	1692
Qy	1778	ggaggtgacagctcttctgacctgagctgcctcgtcaagcgaacccagagacgcgtccg	1837
Db	1697	agaaacacagccagctgtggctgtccctgcgcagcagcaagtgaacgcgcgtgaa	1752
Qy	1838	taagcgaatccgcgcgcgcgaagaagccctcacaggggagcagtttcaactccagctctc	1897
Db	1757	aggtacctgaagatgtcagaagaagtcataacacaggaacgcgtctcgtgcctcggc	1816
Qy	1998	aggatgaactccgcgcgtgtaaaggctccagacaactctggaacgcgcagga-----	1952
Db	1817	aggtccttattatgactgtgcccacagccggcgctgtgacgcggagcgcagagcc	1872
Qy	1953	---dgatcctgtgagtgggccagggtctatgcagaacacaaatgtcagttgcagggc	2008
Db	1877	gatggagtcctgtaggccggcgacacagctgtggaccagtgctcagctgcgcga	1936
Qy	2009	tggaccctattatgatgagcagagaaocgtcattttatgtccaaatggaacttcca	2068
Db	1937	gggaacgctattacacggcccaagcagagcagctgtgtccatgccagcgggaacttcca	1996
Qy	2065	aaatgggaaggaacaaatgacttgaacctatgccaaagacccaagaatactctggggccct	2126
Db	1997	ggagagagaaggcagctctctgcgaccttggcctggag-----tgatgcacca	2047
Qy	2129	gaagaccacagaagcttggaaatgtctgaatgtggagctcgtgtcaacctgtggaata	2188
Db	2048	cgggcctcttggagccaccaagctaccacagctgtgagtgatgtgccctccatgccaca	2107
Qy	2189	tctgcagatgcttgcacttgcgcagctgtgtgccttggcagcttccagactgaagc	2248
Db	2108	ctctgtagatgggttcaagccctgtcagccatgccacgtggtcactaccaactgaagc	2167
Qy	2249	tgtgtgaactctcgtctccctgtggagaggccttgcacacaaactcaggaggactcc	2308
Db	2168	aggcagcaccctatgcttccctgtggttggggcctcaccaagcatgaaggcccat	2227
Qy	2309	tctcttcaggactgtgaaccaggttcaatgttccactccctggacatttctcaacaccac	2368
Db	2228	tctcttcaagactgtgacacaagctccagctgtcccgaggcactactcaacacacag	2287
Qy	2369	cactcacagatgtattcgttgcaccgtgggaacataccagccttgaatttggaaaaataa	2428
Db	2288	cattcacacgctgtattcgtctgtgccatgggtcctctatcagccgaacttcgtcgaactt	2347
Qy	2429	tgtgtttcttgcaggaaatactaagactgactttgatgtctccacaacataaaccca	2488
Db	2348	ctgcagcgcgttccaggaaacacaagcacagactttgatgtcttaccagtgaggccca	2407
Qy	2489	gtgtaaaaacagaagatgtggagggagcctggagagatttcactgggtcatttgaatccc	2548
Db	2408	atcagaagatcgtcagctgtgggtggggagctgggctggagtttcaactgtattatgacccc	2467

	QY	2549	aactaccaggaattatccacgcaacaggagatgttaagtggacatcacacaaccccc	2560
	DB	2468	cactaccgcgggaactaaccaactggttgtagtctggaactcacaaccccacac	2527
	QY	2605	caagcccgcaactctgatcttggttcctcagatctctcccatagagagagctgtgg	2668
	DB	2538	caaagcagaatccattcttggttacacagatctctcccatcagagatgagtgtgg	2587
	QY	2659	ggactatcttgtagtcgaaacacctctccaatactttctggacaacatacgaaacctg	2726
	DB	2588	ggagctctctgtcatgagaaagaaactatccccatctccattaccacttatgagacctg	2647
	QY	2729	ccajactctcgaacgcccaatcgcttcacactccagggtcaagaagcttgagattcagt	2788
	DB	2648	ccajactctcagcgtcccaatgctctccctgcctgcctctccagggaagctctggatacaatt	2707
	QY	2769	caagtccaatgaagggaacagcgtatgaggtgtccaggtcccatcacgtgacatatgaga	2848
	DB	2708	caasacaagcgaagcccaacagcgcccggtctccagattccctatgttaccatcatgaga	2757
	QY	2849	ggactacacgaactcattttaagacatagttctgagatgagctccaggtctcatgacttgaaa	2908
	DB	2768	ggactatgacagctgttgaagacatttgcagatgcgcggctctatgcctctgaaaa	2827
	QY	2909	ccaatcagaaatcttaaagataagaaaccttatacaggtctgtttgagtccctggacca	2969
	DB	2828	ccaacagagagattttaaagcaagaagctctatcaaggctcttttggaggtctagacca	2887
	QY	3026	tcccagactatttcaagtaaacagcccaggagtcocagagagatctccaaagtctgt	3026
	DB	2888	ccccagaactaactcaagtacacag---agaacaacagggagatgcgcgcaaaatctct	2944
	QY	3029	carccvgatgtacgttccaaagtgtccaggtttttgagaccttacaat	3078

25517

RESOL¹,
ABAO3877
ABAO3877 standard; CDNA; 3177 BP.

AC ABA03877;

14-FEB-2002 (first entry)

XX
XXXXX-00000 nucleotide sequence SEQ ID NO:15.
XX

XX Human: POLYX: gamma aminobutyric acid receptor; GABA receptor:
KW epidermal growth factor; EGF: complement receptor; HSP; syntaxin:
KW haemopoietic stem and progenitor cell; sulphotransferase; prohibitin:
KW antiepileptic; cerebroprotective; antiparkinsonian; nootropic; relaxant:
KW anticovulsant; neuroleptic; neuroprotective; antialcoholic; cardiac:
KW tranquiliser; antiarrhythmic; psychiatric; medical; depression; stroke:
KW Parkinson's disease; Huntington's disease; Tourette's syndrome; anxiety:
KW amphoteric lateral sclerosis; head trauma; Alzheimer's disease:
KW alcoholism; vigilance; muscle tension; epileptogenic; memory function:
KW alcoholism; muscle tension; ventricular dysrhythmias.

XX homo sapiens.

[illegible]

ET . CDS 101..3130 /**starv à

```
/product="POLY8"
```

AA
BX
WD260179294-A2.

XX 25-03T-2001.

XX
16-868-2001. 200160-11512854.XX
XXXXX 1000000000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Db 1426 cctgacctgtccctccaggcccccatttttgcacagctctgagatggctccacgttgcg 1485
 QY 1529 ----- 1528
 Db 1486 ctgtgggaacccagcccccagggtgtccacagccctgagctgcccacaaatgggaacagcac 1545
 QY 1529 -----tgtaaccaccatccaggaagctgaacatttcaa 1581
 Db 1546 caactccaaacctgcctgaggtgagtgctccattaaacacagggctccctccaa 1605
 QY 1562 gctaaatgaaggaagctgagttgaa-----aaatgctgagctgtttccagaggtct 1645
 Db 1606 gatcaagatgccaatgcccttttgcactgcgaacaaaggaacaaacagagagggctgg 1665
 QY 1616 gcgcacagactccagagagagcagctcagtaaaagagagctccgcgaactgaact 1675
 Db 1666 agaatcacagggcaggtgtgtccctctgtctgaatgccagctccctccacact 1725
 QY 1676 tacatgcagctgtgcaagcaagctccagagccctgtgcccagacaaagccctcaagga 1735
 Db 1726 taagtgtactctctcagagggcagagggccagagcccgcccgccctccagagcaaga 1785
 QY 1736 aatgttca-----cactgtgtgtgagcttgaaactaaccaagagaggtgcagctc 1782
 Db 1786 ggtcacaaggtccactccctggagctggagcagaggtcagagccgaagaaacacacagccag 1845
 QY 1793 ttgtacactgagctgcactgtaaaagcgaacccagagagcgtccgttaaaagccactccgac 1852
 Db 1846 ctgtgggctgcccctccagacaggaatgggaagcggctgaaagatccctcaaat 1905
 QY 1853 gctcagaagagccgtccacagggagagcttccactccagctcagctcagctgaactcga 1912
 Db 1906 gctcagaagatccataccagggaccctctcctgtgcctgtgcagggcctgtattcga 1965
 QY 1913 ctgtgctaaagacccctcccaaacactctgaaccccgca-----gagctcctgtgg 1963
 Db 1966 gctggtcccaagccgggtgtgtagccggggagcagcagagccgtatggagctgtgag 2025
 QY 1964 agtggggcagggctcagcaaaaaacaaatgtgtcactgtgcaggtctgggacctattcga 2023
 Db 2026 gccggggcagccctgtggtgggaacagtggtcagctgctggcggcgggaatcattacca 2085
 QY 2024 tggagcaggaacgtctcattttatgtccaaatggaaaccttccaaatgaggaagagca 2083
 Db 2086 cggccagacggagcaggtgtgtccatgccagcgggacccctccagagagagagggca 2145
 QY 2084 aatgactgtgaacctgcccaagacaggaatctggggccctggaagcccccagagagc 2143
 Db 2146 gctcctgtgcacttggccctggag-----tgatgccacgggctctctggagc 2186
 QY 2144 ttggaatatgtgaatgtgaggtgtgtcactgtggaactgtggaatattctgcagatggct 2203
 Db 2197 caccacgtcaacacgtgtgaggtgagtgccacactgcccacactctgagatgggtt 2256
 QY 2204 tgcacttgcagctgtgcccctggcagcttccagctgaactgaactgtgtgaacttccctg 2253
 Db 2257 caagccctgtcagccctgcacagctggcactccacactccacactgaagcagagagccctatg 2316
 QY 2264 ctccctctgtggagggagccttgcacacacacacacacagggaggtactcttccctcagagct 2323
 Db 2317 ctccctctgtgtgggggctccacacacacacacagggagggccattctccacagagc 2376
 QY 2324 tgaacacagatccaatgtcactgtgagactttctcactacacacacacacacacaggtat 2383
 Db 2377 tgaacacagatccaatgtcactgtgagactttctcactacacacacacacacaggtat 2436
 QY 2384 tgggtccaggtgggaac 2443
 Db 2437 tgggtgtgcaatgggtctcactacagccgaacttccctgaagaactctggaagcgtgtcc 2496
 QY 2444 aggaataactacactgaactgtgtgagcttgcacacacacacacacacacacacacacacag 2503
 Db 2497 aggaacacacagcagactttgtggtctctaccagtggtggccaatgcaagaatcgtca 2556

RESUL1 8
 ABA0374

ID ABA0374 standard: cDNA: 4440 bp.
 XX ABA0374:

XX ABA0374:
 XX ABA0374:

XX 14-FEB-2002 (first entry)
 Human PLY5 nucleotide sequence SEQ ID NO:9.

Human: PLY5, gamma aminobutyric acid receptor, GABA receptor;
 epidermal growth factor; EGF; complement receptor; HSP; syntaxin;
 haemopoietic stem and progenitor cell; sulphotransferase; prothitin;
 antidepressant; neuroprotective; antiparkinsonian; neurotic; relaxant;
 anticonvulsant; neuroleptic; neuroprotective; antialcoholic; cardiac;
 tranquiliser; antiarrhythmic; psychiatric; medical; depression; stroke;
 Parkinson's disease; Huntington's disease; Tourette's syndrome; anxiety;
 amyotrophic lateral sclerosis; head trauma; Alzheimer's disease;
 alcoholism; vigilance; muscle tension; epileptogenic; memory function;
 cardiomyopathy; arrhythmogenic right ventricular dysplasia; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 679..2886
 /tag= a
 /product= "PLY5"

WC2000175294-A2.
 25-OCT-2001.

otic; antiasthmatic;
sive; hypertensive;
s.

Db 978 agacagagtgcttcttgatcgaaactgtgaccacatattgtccaacacacaggaagctt 1037
QY 1235 tgccttgcttgcaacagagggtacacccctgtatgtcttcacccactgtggaagacacaa 1294
Db 1038 ccagtgctctgcacatcggtgctacactgtgtatgtgtatcaccactgtggtgatgtga 1097
QY 1295 tgaatgacagatcaacaacagaggtgtcagcaggtctgtgtgaacacagtggtgagacta 1354
Db 1098 tgaatgacagatcaacaacagaggtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1157
QY 1355 tgaatgacagatcaacaacagaggtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1411
Db 1158 ccagtgacactgtccacagaggtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1217
QY 1412 agtgaaggggtctctgccacaaagtgtgtcacccgtgtgtccctgcactgtgtgaagag 1471
Db 1218 gccactgaagtgtcaggtgtcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1277
QY 1472 tggtagagagagaggtgtcttctcagatgtcactctgtgtgtgtgtgtgtgtgtgtgtgt 1531
Db 1278 tggcaagaagagacactgt 1333
QY 1532 caccacactcaggaaggtgttaaccttaagcttaagtgaaggaaggtgtgtgtgtgtgt 1591
Db 1334 -----agtgtgagaat 1344
QY 1592 tgcctgagctgttcccgaggt 1651
Db 1345 ggtctcagctgtgagctgt 1378
QY 1652 agagagcttcgt 1711
Db 1379 -----gggt 1429
QY 1712 tggcgaacagacacccctaaagaaatgtttatcactgtgtgtgtgtgtgtgtgtgtgtgt 1771
Db 1430 actgtcagtgaggt 1489
QY 1772 ccaaaagaggt 1831
Db 1490 ccaaaagcgt 1510
QY 1832 gctcgttaagccatcgtcagctcagaagggcgtgtccacagggagaggtgtgtgtgtgt 1891
Db 1511 -----gaacaaaggcaaaacagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1548
QY 1892 gctcagagcatgaacctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1951
Db 1549 gggccaggt 1604
QY 1952 agagtcctgt 2011
Db 1605 tgaactctctgt 1663
QY 2012 gacctattatgt 2071
Db 1664 caaggtcaccctgt 1719
QY 2072 tgagggaagacaaatgactgt 2131
Db 1720 -----ggcctccacacacagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1756
QY 2132 gaccocagaagctgt 2191
Db 1757 tccgacagagtggaacggcgt 1816
QY 2192 tgcagatggctgt 2251
Db 1817 accagagccgt 1876
QY 2252 tcgaacttctgt 2311

Db 1277 gacctgtgagccgt 1936
QY 2312 cttttagaactgt 2371
Db 1937 ct-----gggacaaagt 1998
QY 2372 tcaacagatgtattgt 2431
Db 1969 ccaacgt 2448
QY 2432 tgttttgcgcagaaataactacagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2491
Db 2049 cagcgt 2108
QY 2492 taaacacagagatgt 2551
Db 2109 caaghaatgt 2168
QY 2552 ctaccagaggaattaccac 2611
Db 2169 ctaccagaggaattaccac 2228
QY 2612 ggcgcgcacactgt 2671
Db 2229 gcacagatcctgt 2288
QY 2672 ctatctgt 2731
Db 2289 cgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2348
QY 2732 gacctcagaaacgt 2791
Db 2349 gacctcagaaacgt 2408
QY 2792 gtcacatgaaggaacagcgt 2851
Db 2409 gacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2468
QY 2852 ctaccaggaactcatttaagacatagttcagaggtgtgtgtgtgtgtgtgtgtgtgtgt 2911
Db 2469 ctatgtcagcgt 2528
QY 2912 tcaagaaatacctcagatagaacacacacacacacacacacacacacacacacacacac 2971
Db 2529 ccagagagatttttaagacacagcagcagcagcagcagcagcagcagcagcagcagc 2588
QY 2972 ccagaaactatttcaagacacagcagcagcagcagcagcagcagcagcagcagcagc 3031
Db 2589 ccagaaactatttcaagacacagcagcagcagcagcagcagcagcagcagcagcagc 2645
QY 3032 ccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3078
Db 2646 ccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2692

RESULT 15

AAS70496

ID AAS70496 standard; cDNA; 2723 bp.

XX AAS70496;

XX AAS70496;

CT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #300.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX Food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX W0200175067-A2.

XX 11-OCT-2001.

us-09-747-371-1.rng

Wed Jul 3 08:56:42 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 16:53:19 ; Search time 4:05.21 Seconds
(without alignments)
17435.631 Million cell updates/sec

Title: US-09-747-371-1

Perfect score: 3737

Sequence: 1 ggcgcgcgcgcacacccctcc.....ctaaagcactctcgagacat 3737

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3555312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: gb_vl.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_mu.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_hvg_hum.*

31: em_hvg_inv.*

32: em_hvg_inv.*

33: em_hvg_inv.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	3737	17435.631	1	US-09-747-371-1

C	1	1069.2	48.6	4821	6	AX350811	AX350811 Sequence	
	2	1067.2	48.6	4550	6	AX350813	AX350813 Sequence	
	3	959.4	45.7	3177	6	AX350815	AX350815 Sequence	
	4	940.6	45.2	4440	6	AX350809	AX350809 Sequence	
	5	830.4	41.8	5741	10	AF276425	AF276425 Mus muscu	
	6	813.4	41.6	256164	9	HS4006877	HS4006877 Homo sapi	
	7	754.6	41.0	811	6	AX067305	AX067305 Sequence	
	8	484	48.4	13.0	484	9	HOM2D73H01	AF086390 Homo sapi
C	9	482.8	12.3	65263	2	AC079296	AC079296 Homo sapi	
C	10	364.4	10.3	201805	2	AL606466	AL606466 Mus muscu	
C	11	291	7.5	281	6	AX332310	AX332310 Sequence	
C	12	281	7.5	281	6	AX332310	AX332310 Sequence	
C	13	268.2	7.2	71465	6	AX337325	AX337325 Sequence	
C	14	257	6.9	291	6	AX302926	AX302926 Sequence	
	15	253.6	6.8	234169	10	MM0400878	AX100678 Mus muscu	
	16	213.4	5.7	241	6	AX106631	AX106631 Sequence	
	17	213.4	5.7	241	6	AX140922	AX140922 Sequence	
	18	213.4	5.7	241	6	AX200782	AX200782 Sequence	
	19	213.4	5.7	241	6	AX267438	AX267438 Sequence	
C	20	171.2	4.6	56919	2	AC100761	AC100761 Homo sapi	
C	21	157.2	4.1	199203	2	AC026077	AC026077 Homo sapi	
	22	153	4.1	56919	2	AC100761	AC100761 Homo sapi	
	23	147	3.9	98110	2	AC094770	AC094770 Rattus no	
	24	129.6	3.5	199203	2	AC026077	AC026077 Homo sapi	
C	25	124	3.4	126	11	G59675	G59675 SH3C-130249	
C	26	125	3.3	126	11	G43722	G43722 WIAF-3297-S	
C	27	121.6	3.3	117026	9	HS32945	297632 Human DNA S	
C	28	118.6	3.2	136774	2	AC096475	AC096475 Rattus no	
C	29	118.2	3.2	129389	9	HS26114	282214 Human DNA S	
C	30	113	3.1	175192	2	AC044848	AC044848 Mus muscu	
C	31	113	3.1	220050	10	AL583887	AL583887 Mouse DNA	
C	32	111.6	3.0	2465	6	A31598	A31598 M.musculus	
C	33	111.6	3.0	2465	10	MMURPER	X02965 Mouse DNA S	
C	34	111.6	3.0	2465	10	MMURPER	X02965 Mouse DNA h	
C	35	109.6	2.9	98110	2	AC094770	AC094770 Rattus no	
C	36	98.4	2.6	136774	2	AC096475	AC096475 Rattus no	
	37	96.4	2.6	832	11	G72521	G72521 MARC 2843-2	
	38	89	2.4	3969	10	BC005443	BC005443 Mus muscu	
C	39	88.4	2.4	635	11	G72504	G72504 MARC 2901-2	
C	40	86	2.4	81961	9	HS100422	295756 Human DNA S	
	41	87.4	2.3	1663	6	AX023978	AX023978 Sequence	
	42	87.4	2.3	4377	10	MMURPER	X5285 M.musculus	
C	43	86.6	2.3	217078	2	AL591964	AL591964 Mus muscu	
	44	84.6	2.3	2746	9	HS4602406	AL137638 Homo sapi	
	45	84.6	2.3	2863	6	AX015418	AX015418 Sequence	

ALIGNMENTS

RESULT 1
AX350811/c
LOCUS AX350811 4821 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 11 from Patent WO0175294.
ACCESSION AX350811
VERSION AX350811.1 GI:18616291
KEYWORDS unclassified.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (sites)
AUTHORS Taglier, R.J., Vernet, C.A., Fernandes, E., Shinkets, R.A.,
Majumder, K., Padigara, M., Colman, S.D., Zerkow, B.D., Seytek, K.A.,
Bargess, C.E., and Liu, X.
TITLE Novel human proteins, polynucleotides encoding them and methods of
using the same
JOURNAL Patent: WO 0175294-A 11 25-OCT-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..4821
/organism="unclassified"
/db_xref="taxon:32644"
/note="POLYX 297832_B.0.707"

QY 1016 cactctccagcttgatggaagacatgtataagattatgagtgccagacccgcaatg 1075
DB 917 CATGCTGAGCAGACAGAGAGAGCGTGCACAAATATAGATAGTGGCTTAACACAGG 976
QY 1076 agtttgatcattcttgcacaaacatctgtggcgagttttgactgagtggtgcacaaag 1135
DB 977 GGCCTGTGACCATATTTCCGCCAACACAGTGGGCAGCTTCGAATGCCAGTTGCAAGAAAG 1036
QY 1136 acttaaatattacagataagaagcttgcacaaatgtgagatgtgagtgctcttggatag 1195
DB 1037 CTATAGCTTCATCAATAGAGAGAGTCCAGGATATAGACAGTGTCTTTGATCG 1086
QY 1136 gactgtgaccacagctgcatacaacacccctgtgcacatttctgtgtcttgcaacagag 1255
DB 1097 AACCTGTGACCATATGTGTCAACACACACAGAGAGCTTCCAGTGTCTCTGCAATGCTGG 1156
QY 1256 gtacacccctgtatggtctcaccacactgtggagacccaatagtgagcagcatcaacaaag 1315
DB 1157 CTACTGTGTATGATATCAACCCACTGTGGGATGTGGATGAATGCAGCATCAACCGGGG 1216
QY 1316 aggtgtcagcaggtctgtgtgacacagtgggcagcttatgaaagccagtgccacccctgg 1375
DB 1217 AGGTGTGCGCTTGTGCTCATCAACTCTCTGSCAGCTAACCAGTGTACTCTGCCAGCAGG 1276
QY 1376 --gtacacgtccactggataaaagaactgtgtgaaagtgaagggctctctgccac 1432
DB 1277 CCAGGTGTGCTGTCACTGGAATGCAAAAGATTCACAGAGCCACTGAAGTGTGAGGGGAG 1336
QY 1433 aagtgtgtcacccctgtgtctctgcaactgcggtgaagtggtgtgagagagaggtgtct 1492
DB 1337 TCCGTGGGCTCGAAAGCCAGTGTCACTGTCAACCGGTCTGGCAAGAGACACCTGTGC 1396
QY 1493 octcagatgtcactgtgcatcactctctctctcactctcactcactcactcactcactc 1546
DB 1397 OCTGAOCTGTCTCCAGGGCCCGATTTTGGCCAGAGGCTCAGTGTCTCTCTTAACA 1456
QY 1547 aagtgttaaccttaagttaagtgaagcaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1600
DB 1457 ACGGGCTCTCTCAAGATCAAGGATGCCAANTGCCGTGTTCACCTGGCAACAAAGGAA 1516
QY 1601 gttcccgaggggtctgtgcacagcactctacagagagcagctcagtaaaagagagctt 1660
DB 1517 AACAGAGAGGGTGTGCACAAACAGAGGGCGAGTGGTGGCCCTCTCTCTGTAATGCAGGT 1576
QY 1661 ccgtgaactaaaccttactcagctctgtgcaagcaagtctccagagagccctggtccgacc 1720
DB 1577 CACTTCTACCTTATAGTGTGACTCTCTCGGAGGGGCAAGGCGCGAGGCGCGCGAC 1636
QY 1721 aagcacccttaagaaatgttta---tcaactgttggatttgagcttgaaactaaccataa 1777
DB 1637 CCTCCAGGCAAGAGGTGCACAGGCTCACCTGTGAACCTGGAGGTCAGAGTCAAGGCA 1696
QY 1778 ggaagtgacagctcttctgtgacctgagctgcatgttaagcgaaacgagaaagcgtccg 1837
DB 1637 AGAACACAGCAGCAGCTGTGGCTGCTCCCTCCAGCAGCAATGTGAAGCGGCTGAA 1756
QY 1838 taagaccaatccgacagctgcacaaagggcgctccacagggagcagtgcttcaactcagctctc 1897
DB 1757 AGGATCCTGAAGATGCTCAAGAAAGTCCATCAACAGGAGCGCTTCTGTGTGCGCGCTG 1816
QY 1898 aggcataagcctcagctggtctaaagagctcccgagaactctgaagccaggtca---- 1952
DB 1817 AGGCTTGATATAGCTGTGGCCCAAGCGCGGCTGTGTAGCCGGGAGGAGCAGAGCC 1876
QY 1953 ----gagctctgtggagtgggcaggtctcagtcagaaacacaaatgtcagtrtggagggc 2008
DB 1877 GATGAGTCTGTAGGCCCGGCGAGCACCTGTGTGGACCAAGTGTGTCACTGTCCGCCA 1936
QY 2009 tggagacattatgtgagcaagcagcgtgacttttatgtccaaatggaaacttcca 2058
DB 1937 GGGAAAGCTATTACAGGCCAGAGCGAGCAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1996
QY 2069 aaatgaggaaggaataatgactgtgaaacctgtgcaacagcaggaataatctgtggcgct 2126

DB 1997 GGAAG 2047
QY 2129 gaaagaccccaagctgtgaaatgtctgaaatgtgaaatgtgaaatgtgaaatgtgaaatgt 2188
DB 2048 GGGGCT 2107
QY 2189 ttcctcagagaggttttgcacactgtgcaagctctgtgcccgtgtgcccgtgtgcccgtgt 2248
DB 2108 CAGTGTAGT 2167
QY 2249 tggctgaacttctgtcttccctgtggaggggagccctgtgcccacacacacacacacac 2308
DB 2168 AGGACGGAGCTGTGCTTCT 2227
QY 2309 ttccttccagagaggtgtgaaacagaggttcaagtctcaactgtgacatttctcaacacac 2368
DB 2228 TCCCTTCCAAAGT 2287
QY 2369 cactcaccagatgtatttctgttgcacagtggaacataccagcctgaaatttggaaaaataa 2428
DB 2288 CATCCAGCGT 2347
QY 2429 tttgttttcttccagagaaactaactgaactgtgacttggactgtgactgtgactgtgact 2488
DB 2348 CAGGAGCGGCTGT 2407
QY 2489 gtttaaaag 2548
DB 2408 AAG 2467
QY 2549 aagatccacag 2608
DB 2468 CAGCTTACCGGCGAG 2527
QY 2609 caagcgccagcactctgt 2668
DB 2528 CAGGCGCAGAGTCTGT 2587
QY 2669 ggaactatctgt 2728
DB 2588 GGAAGCTCTGT 2647
QY 2729 ccagactcag 2788
DB 2648 CAGAGCTTACAG 2707
QY 2789 caagtccaatggaag 2848
DB 2708 CAAGACAG 2767
QY 2849 ggaatccag 2908
DB 2768 GGAATATGAG 2827
QY 2909 ccactcag 2968
DB 2828 CAG 2887
QY 2969 tccacag 3028
DB 2888 CAG 2947
QY 3029 cagctgattgtgacttccaaag 3078
DB 2945 CAG 2994

RESULT 3
AX350815
LOCUS
DEFINITION Sequence 15 from Patent WO0179294.
ACCESSION AX350815

QY 1030 atggaagacatgataatgagtgccagagccagcccaatgaggttgatcatt 1089
 DB 827 atggaagacatgcaaaagacatcaccagtgcttgatgaatnagcaggtggacacat 886
 QY 1090 tctgcaaaaacatctggggcagttttgactggcggtgcaagaagatttaattattaa 1149
 DB 887 tctgcccacacacccgtggcagctttgaattggcgttcccaaaaggccacacacactga 946
 QY 1150 cagatgagaagctcttccaagatggtgatgagtgctcttggatagagactgtgacaca 1209
 DB 947 cggagcggcggacatcccaagacatgacagaggttccctttgagcggcgttgacaca 1006
 QY 1210 gctgcataaccacccctggacatttctgtgtgtgtgcaaccaggtgacacctgtatg 1269
 DB 1007 tctgcatcaactcccccggcagtttccagtgccctgtgtgcccggcaggttacacacactga 1066
 QY 1270 gcttcaaccacttgagacacacacatgagtgagagctcaaacagagagctgacagcag 1329
 DB 1067 ggaacacccactcgggagacgtgggatgaatgtacatgaacacacagcagcttgacagcag 1126
 QY 1330 tctgtgtgaacacagtgggcagctatgaatgccagtgccacccctgggtgacaagctccact 1389
 DB 1127 gctgtgtcaacacacagggcagctgaaggtgtgtgtgcccacacagggagcgtgacat 1186
 QY 1390 ggaataaaaagactgtgtggaagtgaagggctcctg---cccaaggtgtgtcacccc 1446
 DB 1187 ggaatcagaagcactgtgtggagatgaacgggtgctgtctggtcgaagcgtctgccc 1246
 QY 1447 gctgtcctcgcactcggtaagtggtgtgagagagcgggtgctcctcagatgtcact 1506
 DB 1247 agcccccagctgtcctgtgcccaggtgggtggagtggaagtggaacgtgtctgtcctgctg 1306
 QY 1507 ctggttaccctc-----tcttgaatgtccaccacctcaggaagtgta----- 1553
 DB 1307 gccacagctcttctcagccagattcagaagaacagcactacactcctgagctgtgtgtccag 1366
 QY 1554 ----- 1553
 DB 1367 gtctccaggcagacactgcccacacacatggcactagctccagtagcagggcccggt 1426
 QY 1554 -----accttaagataaatgaagcgaagtgtgtttgaaaatgtgt 1596
 DB 1427 gctcagatgccccaccacccctatcagacagagagggccggttcacagatgacca 1486
 QY 1597 agctgttctccagaggtgtggaacagcaactacacagaa-----gcacagctcagtaa 1650
 DB 1487 agtgcctatctccagccccaccacagcagcagcagcagcagcagcagcagcagcagcagc 1546
 QY 1651 aagagagcttccgtacgtacaaaccttacatgcagctctggaagcaagctccagggagccc 1710
 DB 1547 tagacactgcccagctgactttgtgacccctcaagtgactgactcctcccaagagagcggcc 1606
 QY 1711 ctggcccaacacacccctcaggaatgt---ttacactgttgagttgagctgaaa 1767
 DB 1607 gaggcccgacgtcccatccaaaggaggtgtccacatcacacagagtttgagctggaga 1666
 QY 1768 ctacccaaaaggaggtgacagctctctgtgacctgagctgcaatgaagcgaacccagaa 1827
 DB 1667 tgaagtgatgagggcttcagccactgcgacagcagactgcaatggcagaggtgcccagc 1726
 QY 1828 agcgtcctgtaagcactccgacgtcagcagcagcagcagcagcagcagcagcagcagcagc 1887
 DB 1727 agagcctcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1786
 QY 1888 tccagctcctcagctgacactcagctgagctgagcgttaaaagcctcccaagaacatctgaagcc 1947
 DB 1787 tccagcttggcagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1846
 QY 1948 aggcagctcctgtgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 2007
 DB 1847 ---ccgggacgttggcagctgagctgagctgagctgagctgagctgagctgagctgagct 1903
 QY 2008 ctgggacacttatgatggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2067

DB 1904 ctgcaatctcttgaatgagtgacccagcagctgagctgagctgagctgagctgagctgagct 1963
 QY 2068 aaatctgtggaagcacaatgactgtggaacccatgcccgaagacagcagcagcagcagcagcagc 2127
 DB 1964 agaaataaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2014
 QY 2128 tgaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2187
 DB 2015 tggcgttggcgttggcgttggcgttggcgttggcgttggcgttggcgttggcgttggcgttggcgtt 2074
 QY 2188 attctgtgagatggtttgtgacactgctgagctgtgtgcccgggagcagcttccagcagcagc 2247
 DB 2075 tctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2134
 QY 2248 ctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2307
 DB 2135 ctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2194
 QY 2308 ctctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2367
 DB 2195 cctcgttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2254
 QY 2368 cctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2427
 DB 2255 cctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2314
 QY 2428 attctgt 2487
 DB 2315 attctgt 2374
 QY 2488 agtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2547
 DB 2375 actcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2434
 QY 2548 caaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2607
 DB 2435 cctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2494
 QY 2608 cctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2667
 DB 2495 caaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2554
 QY 2668 ggaacactcctgtgagctggaacaccccttctcactcactcactcactcactcactcactcactcact 2727
 DB 2555 gtaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2614
 QY 2728 gcaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2787
 DB 2615 gcaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2674
 QY 2788 tcaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2847
 DB 2675 tcaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2734

RESULI 6

HSA160677

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

HSA160677 256164 bp DNA linear PRI 06-FEB-2002
 Homo sapiens ASL3 gene, CENP1 gene, Cllorf14 gene, Cllorf15 gene
 Cllorf16 gene and Cllorf17 gene.

ACCESSION AJ401977.1 GI:8052236
 VERSION ASL3 gene; Cllorf14 gene; Cllorf15 gene; Cllorf16 gene; Cllorf17 gene; CENP1 gene.
 KEYWORDS Homo sapiens

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (sites)
 Amis J., Bahr A., Mujica A., Sampson N., Bikar S.E.,
 Winterpacht A., Zabel B., Hanke H., and Schmidt E.R.
 Comparative genomic sequencing reveals a strikingly similar

architecture of a conserved syntenic region on human chromosome

ilp15.3 (including gene ST5) and mouse chromosome 7

Cytogenet. Cell Genet. 93 (3-4), 284-290 (2001)

21418998

2 (bases 1 to 256164)

Direct Submission

Submitted (27-APR-2000) Hankeln T., Inst. Molekulargenetik, Univ.

Mainz, Becherweg 32, D-55099 Mainz, GERMANY

Location/Qualifiers

1..256164

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="ilp15.3"

/clone="LLNLP704J1411703"

/clone="PAC27561"

/clone_lib="LLNLP704"

18454..18666

/gene="CEGPl"

/number=1

join(18454..18666,20233..20355,30545..30615,35139..35573,

40559..40684,43241..43357,44074..44163,49530..49646,

50629..50751,54145..54261,56293..56415,56837..56953,

59338..59442,61950..62105,62485..62692,76251..76421,

79122..79289,80002..80163,82486..82683,84193..84345,

88092..88171,88850..89659)

/gene="CEGPl"

/evidence="experimental"

18454..89659

/gene="CEGPl"

join(19533..18666,20233..20355,30545..30670,35139..35573,

40559..40684,43241..43357,44074..44163,49530..49646,

50629..50751,54145..54261,56293..56415,56837..56953,

59338..59442,61950..62105,62485..62692,76251..76421,

79122..79289,80002..80163,82486..82683,84193..84345,

88092..88171,88850..89002)

/gene="CEGPl"

/codon_start=1

/product="CEGPl protein"

/protein_id="CAB92285.1"

/db_xref="GI:8052237"

/db_xref="SPTREMBL:Q9NQ36"

/translation="MGVAGNRKPCAAVLLILLPLILLAGVPPGRRAGPQP

DVDCAGLDGDHADALQNTPTSYKCSKPCVQGRQEDIDSCNELNGCVHDC

LNIPGYKTCDEGPMALHDHGNCLDVDECLNGCCGCTVNVVNGYECCECFLL

SDNGHTCIHRSSEGLSCMKNKHGCHICKREAPGVSACECPGFLAKNORDILTCN

HNGGCGHDDTDADGPCSCPKYKMTDGRSLEREDIVLEVTESNITISVYGGDKR

VKRLMETCAVNNGCDDTKDSTGVHSCFPGFTLLQDGTCKDIDECQTRNGC

DHFCNIVSFDQCKGKGLLTDEKSCQDVECLDRTCDHSCINHPGFACQNRG

YTLGPTHCQDTEGCSINNGCCQCVNTVGSVQCQHPGYKLHNNKDCVVKGLLP

TSVSPVSLHCKSGGGQCFLRCHGHLSSDVITLTSVTFKLESGKSLKAEFL

PEGLRALPKHSSVKESFRVNLTCSSKQVPGAPGAPSTPKEMFIIVEFELEINQK

EVTASDLSICVIRTEKRLKRAIRLRKAREQFLQLSGNLDVAKPRISERQA

ESCGGQHAECVSCRAGTYDGAERCLCPNGITFQNEEGQTCPCPCPNCSGA

LKTPAANNSECGGCGEYSGADGFCALGTFFPEAGRTSCPCGGSLAKHQ

GATFQDCQTRVQCSPGHFYNTTHRCIRCPVGTQVPEFGKNCVSCPGNITIDFDS

PNITQCKNRRCGELGDFGTGIESPNVPGNPANTECTWIPPPKRIILIVPEFL

PEDDCCDILVWRKTSNNTVYETQCTVERPIATFSKSKLNIOKSKNENSGAEF

QVPVYVDEYDELLIDVRGRLYASENHQELIKKLLKALFDVLAHFQNFKYIA

QESREMPRSFIRLSKSVSRFLPYK"

18667..20232

/gene="CEGPl"

/number=1

20233..20355

/gene="CEGPl"

/number=2

20356..30544

/gene="CEGPl"

/number=2

30545..30670

/gene="CEGPl"

/number=3

intron

exon

intron

exon

intron

/gene="CEGPl"

/number=3

35139..35573

/gene="CEGPl"

/number=4

35574..40558

/gene="CEGPl"

/number=4

40559..40684

/gene="CEGPl"

/number=5

40685..43240

/gene="CEGPl"

/number=5

43241..43357

/gene="CEGPl"

/number=6

43358..44073

/gene="CEGPl"

/number=6

44074..44163

/gene="CEGPl"

/number=7

44164..49529

/gene="CEGPl"

/number=7

49530..49646

/gene="CEGPl"

/number=8

49647..50628

/gene="CEGPl"

/number=8

50629..50751

/gene="CEGPl"

/number=9

50752..54144

/gene="CEGPl"

/number=9

54145..54261

/gene="CEGPl"

/number=10

54262..56292

/gene="CEGPl"

/number=10

56293..56415

/gene="CEGPl"

/number=11

56416..56836

/gene="CEGPl"

/number=11

56837..56953

/gene="CEGPl"

/number=12

56954..59337

/gene="CEGPl"

/number=12

59338..59442

/gene="CEGPl"

/number=13

59443..61949

/gene="CEGPl"

/number=13

61950..62105

/gene="CEGPl"

/number=14

62106..62485

/gene="CEGPl"

/number=14

62486..62692

/gene="CEGPl"

/number=15

62693..76250

/gene="CEGPl"

/number=15

intron

KEYWORDS HTG: HTGS_PHASE0.
SOURCE Human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 65263)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-467K18
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 65263)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Bedalov, J., Boguski, M.,
 Bouckgeater, B., Brown, A., Burkett, G., Campoliano, A., Castle, A.,
 Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kam, L., Karas, A., Lakocque, K.,
 Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Lieu, C., Liu, G.,
 Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, H., Riley, R.,
 Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theadore,
 Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Yi, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zimmer, A. and Zody, M.
Direct Submission
TITLE Submitted (26-AUG-2000) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center -----
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: W18K
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: LL0929
 Center clone name: 467_K_18

 * NOTE: This record contains 79 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be generic and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 1
 733 832: contig of 732 bp in length
 833 1574: contig of 742 bp in length
 1575 1674: gap of 100 bp
 1675 2417: contig of 743 bp in length
 2418 2517: gap of 100 bp
 2518 3233: contig of 716 bp in length
 3234 3333: gap of 100 bp
 3334 4028: contig of 695 bp in length
 4029 4128: gap of 100 bp
 4129 4821: contig of 693 bp in length
 4822 4921: gap of 100 bp
 4922 5653: contig of 732 bp in length
 5654 5753: gap of 100 bp
 5754 6479: contig of 726 bp in length
 6480 6579: gap of 100 bp
 6580 7299: contig of 720 bp in length
 730 7399: gap of 100 bp
 740 8128: contig of 729 bp in length
 8129 8226: gap of 100 bp
 8227 8542: contig of 714 bp in length
 8543 9042: gap of 100 bp
 9043 9752: contig of 750 bp in length
 9753 9892: gap of 100 bp
 9893 10631: contig of 739 bp in length
 10632 10731: gap of 100 bp
 10732 11451: contig of 720 bp in length
 11452 11551: gap of 100 bp
 11552 12286: contig of 735 bp in length
 12287 12366: gap of 100 bp
 12367 13132: contig of 746 bp in length
 13133 13232: gap of 100 bp
 13233 13982: contig of 750 bp in length
 13983 14062: gap of 100 bp
 14063 14836: contig of 754 bp in length
 14837 14936: gap of 100 bp
 14937 15665: contig of 729 bp in length
 15666 15765: gap of 100 bp
 15766 16483: contig of 718 bp in length
 16484 16583: gap of 100 bp
 16584 17289: contig of 706 bp in length
 17290 17389: gap of 100 bp
 17390 18090: contig of 701 bp in length
 18091 18190: gap of 100 bp
 18191 18907: contig of 717 bp in length
 18908 19007: gap of 100 bp
 19009 19726: contig of 719 bp in length
 19727 19826: gap of 100 bp
 19827 20541: contig of 715 bp in length
 20542 20641: gap of 100 bp
 20642 21421: contig of 780 bp in length
 21422 21521: gap of 100 bp
 21522 22226: contig of 705 bp in length
 22227 22326: gap of 100 bp
 22327 23076: contig of 750 bp in length
 23077 23176: gap of 100 bp
 23177 23909: contig of 733 bp in length
 23910 24009: gap of 100 bp
 24010 24755: contig of 746 bp in length
 24756 24855: gap of 100 bp
 24856 25573: contig of 718 bp in length
 25574 25673: gap of 100 bp
 25674 26396: contig of 723 bp in length
 26397 26496: gap of 100 bp
 26497 27234: contig of 738 bp in length
 27235 27334: gap of 100 bp
 27335 28079: contig of 745 bp in length
 28080 28179: gap of 100 bp
 28180 28872: contig of 693 bp in length
 28873 28972: gap of 100 bp
 28973 29727: contig of 755 bp in length
 29728 29827: gap of 100 bp
 29828 30583: contig of 756 bp in length
 30584 30683: gap of 100 bp
 30684 31431: contig of 748 bp in length
 31432 31531: gap of 100 bp
 31532 32207: contig of 676 bp in length
 32208 32307: gap of 100 bp
 32308 33033: contig of 726 bp in length
 33034 33133: gap of 100 bp
 33134 33814: contig of 681 bp in length
 33815 33914: gap of 100 bp
 33915 34636: contig of 722 bp in length
 34637 34736: gap of 100 bp
 34737 35456: contig of 720 bp in length
 35457 35556: gap of 100 bp
 35557 36279: contig of 723 bp in length
 36280 36379: gap of 100 bp
 36380 37113: contig of 734 bp in length
 37114 37213: gap of 100 bp

Db 261 ATGTCCTGGCCATCCCGACAGATATTTCAGTACACAGCCCGAGAGTCCCGAGAGATJ 222
 QY 3016 ttccagagatgcttcacgattgctacgttcacaaagtccacagtttccagttttgagacattaca 3075
 Db 221 TTCAGAGTCGTTCATCGATGCTACGTTCCAAAGTTCACAGTTCACAGTTCACAGTTCAC 162
 QY 3076 aatgaactcagccacgctcccaactcaatacaaatgtctctgctatagagttggtggtgagacaga 3135
 Db 161 AATGACTCAGCCAGCTCCGACTCANTACAAATGTTCTGCTATAGAGTTCGTGGGACAGA 102
 QY 3136 gctgtctctctctgctacgttcagcaagtcgggttattgctgctccctccatcagtcgactc 3195
 Db 101 GCTGCT 42
 QY 3196 attagagttcaatttttatagataatacacagatatttsgta 3236
 Db 41 ATTAGAGTTCAMTTTATAGTAATACAGATATTITGTA 1
 RESULT 13
 AC101119 71465 bp DNA linear HIG 23-NOV-2001
 LOCUS Mus musculus clone RP23-93F3, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION AC101119
 ACCESSION AC101119
 VERSION AC101119.1 GI:17059893
 KEYWORDS HTG; HTGS-PHASE0.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 71465)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-93F3
 Unpublished
 2 (bases 1 to 71465)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.S., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kanat,A., Karatas,A., Kellis,C., Laboche,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheters,K., Melgrim,J.,
 Meneses,L., Minova,T., Mienga,V., Murphy,I., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,S.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L14980
 Center clone name: 91_F_3

 * NOTE: This record contains 88 individual

* sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 726: contig of 720 bp in length
 721 820: gap of 100 bp
 821 1643: contig of 723 bp in length
 1644 1643: gap of 100 bp
 1644 2355: contig of 712 bp in length
 2356 2455: gap of 100 bp
 2456 3151: contig of 696 bp in length
 3152 3251: gap of 100 bp
 3252 3985: contig of 734 bp in length
 3986 4085: gap of 100 bp
 4086 4815: contig of 730 bp in length
 4816 4915: gap of 100 bp
 4916 5648: contig of 733 bp in length
 5649 5748: gap of 100 bp
 5749 6466: contig of 718 bp in length
 6467 6566: gap of 100 bp
 6567 7284: contig of 718 bp in length
 7285 7384: gap of 100 bp
 7385 8057: contig of 713 bp in length
 8058 8157: gap of 100 bp
 8158 8915: contig of 718 bp in length
 8916 9015: gap of 100 bp
 9016 9749: contig of 734 bp in length
 9750 9849: gap of 100 bp
 9850 10571: contig of 722 bp in length
 10572 10671: gap of 100 bp
 10672 11375: contig of 704 bp in length
 11376 11475: gap of 100 bp
 11476 12188: contig of 713 bp in length
 12189 12888: gap of 100 bp
 12889 12988: contig of 700 bp in length
 12989 13088: gap of 100 bp
 13089 13811: contig of 723 bp in length
 13812 13911: gap of 100 bp
 13912 14627: contig of 716 bp in length
 14628 14727: gap of 100 bp
 14728 15418: contig of 691 bp in length
 15419 15518: gap of 100 bp
 15519 16225: contig of 707 bp in length
 16226 16325: gap of 100 bp
 16326 17028: contig of 703 bp in length
 17029 17128: gap of 100 bp
 17129 17346: contig of 720 bp in length
 17347 17438: gap of 100 bp
 17439 18557: contig of 709 bp in length
 18558 18757: gap of 100 bp
 18758 19463: contig of 706 bp in length
 19464 19563: gap of 100 bp
 19564 20268: contig of 705 bp in length
 20269 20368: gap of 100 bp
 20369 21067: contig of 719 bp in length
 21068 21167: gap of 100 bp
 21168 21910: contig of 723 bp in length
 21911 22010: gap of 100 bp
 22011 22729: contig of 719 bp in length
 22730 22829: gap of 100 bp
 22830 23541: contig of 712 bp in length
 23542 23641: gap of 100 bp
 23642 24348: contig of 707 bp in length
 24349 24448: gap of 100 bp
 24449 25150: contig of 702 bp in length
 25151 25250: gap of 100 bp
 25251 25946: contig of 696 bp in length

```

* 25847 26046: contig of 100 bp
* 26047 26774: contig of 728 bp in length
* 26775 26874: gap of 100 bp
* 26875 27581: contig of 707 bp in length
* 27582 27681: gap of 100 bp
* 27682 28379: contig of 698 bp in length
* 28380 28479: gap of 100 bp
* 28480 29184: contig of 705 bp in length
* 29185 29284: gap of 100 bp
* 29285 30009: contig of 725 bp in length
* 30010 30109: gap of 100 bp
* 30110 30825: contig of 716 bp in length
* 30826 30925: gap of 100 bp
* 30926 31640: contig of 715 bp in length
* 31641 31740: gap of 100 bp
* 31741 32453: contig of 713 bp in length
* 32454 32553: gap of 100 bp
* 32554 33271: contig of 718 bp in length
* 33272 33371: gap of 100 bp
* 33372 34065: contig of 694 bp in length
* 34066 34165: gap of 100 bp
* 34166 34867: contig of 702 bp in length
* 34868 34967: gap of 100 bp
* 34968 35674: contig of 707 bp in length
* 35675 35774: gap of 100 bp
* 35775 36490: contig of 716 bp in length
* 36491 36590: gap of 100 bp
* 36591 37304: contig of 714 bp in length
* 37305 37404: gap of 100 bp
* 37405 38116: contig of 712 bp in length
* 38117 38216: gap of 100 bp
* 38217 38941: contig of 725 bp in length
* 38942 39041: gap of 100 bp
* 39042 39765: contig of 724 bp in length
* 39766 39865: gap of 100 bp
* 39866 40589: contig of 724 bp in length
* 40590 40689: gap of 100 bp
* 40690 41405: contig of 716 bp in length
* 41406 41505: gap of 100 bp
* 41506 42199: contig of 694 bp in length
* 42200 42299: gap of 100 bp
* 42300 43009: contig of 710 bp in length
* 43010 43109: gap of 100 bp
* 43110 43806: contig of 697 bp in length
* 43807 43906: gap of 100 bp
* 43907 44618: contig of 712 bp in length
* 44619 44718: gap of 100 bp
* 44719 45379: contig of 661 bp in length
* 45380 45479: gap of 100 bp
* 45480 46168: contig of 689 bp in length
* 46169 46268: gap of 100 bp
* 46269 46983: contig of 715 bp in length
* 46984 47083: gap of 100 bp
* 47084 47794: contig of 711 bp in length
* 47795 47894: gap of 100 bp
* 47895 48613: contig of 719 bp in length
* 48614 48713: gap of 100 bp
* 48714 49454: contig of 741 bp in length
* 49455 49554: gap of 100 bp
* 49555 50270: contig of 716 bp in length
* 50271 50370: gap of 100 bp
* 50371 51085: contig of 715 bp in length
* 51086 51185: gap of 100 bp
* 51186 51897: contig of 712 bp in length
* 51898 51997: gap of 100 bp
* 51998 52705: contig of 708 bp in length
* 52706 52805: gap of 100 bp
* 52806 53514: contig of 709 bp in length
* 53515 53614: gap of 100 bp
* 53615 54324: contig of 710 bp in length
* 54325 54424: gap of 100 bp
* 54425 55142: contig of 718 bp in length
* 55143 55242: gap of 100 bp

```

```

* 55243 55962: contig of 720 bp in length
* 55963 56062: gap of 100 bp

Query Match      7.2%; Score 268.2; DB 2; Length 71465;
Best Local Similarity 70.0%; Pred. NO. 5.2e-16;
Matches 376; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

QY 2461 catatgagccttaattgggaaaaataattgttttttccccagggaaaataactacgactg 2460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65957 CCTATAGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65116

QY 2461 acttttattgggtccacaaataacacccagtgtaaaaaacagaagatgtgagggaggtgg 2520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65117 ACTTATAGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65176

QY 2521 gatttttactgggtatcatggaatccccaaactccccaggcaattccccagccaaacacg 2580
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65177 GTGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65236

QY 2581 agtataatggagacatcaacaccccccacacccagccagccatccctgattcctcctaga 2640
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65237 ACTTATAGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65296

QY 2641 tcttttttccctatagaggagcactgtggggactatcttggtgatgcgggaaaacacttctcat 2700
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65297 TCTTATAGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65413

QY 2761 ccaggtt aaataagctgtgatttcaatccaatgaagggagacagcgcctagagsgt 2820
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65414 CCGGTTTAAAGAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65473

QY 2821 tccaggtcccatatgacatargatgaggactccacaggaaactcattgaagacatgctc 2880
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65474 TCCAGTCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65533

QY 2881 gatttttccaggtctatgctatgagaaacacacacacacacacacacacacacacacac 2937
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65534 AGGATCTAGCTTGTGTCAGAGCTTGTGCAACAGTCCGACAGAAATACAGCACTG 65590

RESULT 14
AX302926/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homosapiens
Homo sapiens
REFERENCE
1. (sites)
AUTHORS
Bianchi, F., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and
Kaplan, R. L.
TITLE
Comparisons and methods for the therapy and diagnosis of breast
cancer
JOURNAL
PATENT: W/ 0179286-A 246 25-OCT-2001;
GRIXX CORPORATION (US)
FEATURES
location/Qualifiers
source
1..291
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 70 a 75 c 65 g 71 t 6 others
ORIGIN

Query Match      6.9%; Score 257; DB 6; Length 291;
Best Local Similarity 97.5%; Pred. NO. 1.1e-43;
Matches 268; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

QY 2005 gactggaccattatgatggagcagagaacgc-tgcattttatgtccaaatgaacac 2064
 Db 275 GCGTGGGACCTTATATNATGACGACGAGACGCTTCAATTTATGTCACAAATGGAACT 216
 QY 2065 tccaaatgaggaagacaaatgactgtgaacacatgcccaagaccaggaattctggg 2124
 Db 215 TCCAAATGAGAGAGACCAATGACTGTGNACCATGCCACAGCAGAAATCTGGG 156
 QY 2125 cactgaagaccocccagaagcttggaatatgtctgaattgagtggtgtgtcaacactggtg 2184
 Db 155 CCGTGAAGACCCCAAGAGCTTGAATATGTCTGAATGTGAGGCTGTGTCAACCTGG 96
 QY 2185 aatattctgcagatggcttgacactgccagctctgtgcctggggcagcttccagctg 2244
 Db 95 AATATCTCGAGANGCTTTCACCTGCCAGCTCTGTGCCCTGGCAGCTTCCAGCTG 36
 QY 2245 aagctggctgaactctctctccctctgtggagga 2279
 Db 35 AAGCTGGTCNAACTCTCTCTCCCTCTGGGAGGA 1

RESULT 15
 MMU400878 234169 bp DNA linear ROD-05-FEB-2002
 LOCUS Mus musculus Ascl3 gene, Cegpl gene, D7Hilorf14 gene, D7Hilorf15
 DEFINITION gene, D7Hilorf16 gene and Cllorf17 gene.
 ACCESSION AJ400878
 VERSION 1
 KEYWORDS Ascl3 gene; Cllorf17 gene; Cegpl gene; D7Hilorf14 gene; D7Hilorf15
 gene; D7Hilorf16 gene.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 AUTHORS Amid,C., Hahr,A., Mujica,A., Sampson,N., Bikar,S.E.,
 Winterpacht,A., Zabel,B., Hankeln,T. and Schmidt,E.R.
 TITLE Comparative genomic sequencing reveals a strikingly similar
 architecture of a conserved syntenic region on human chromosome
 1p15.3 (including gene SRS) and mouse chromosome 7
 Cytogenet. Cell Genet. 93 (3-4), 284-290 (2001)
 21418998
 REFERENCE 2 (bases 1 to 234169)
 Hankeln,T.
 TITLE Direct Submission
 JOURNAL Submitted (27-APR-2000) Hankeln T., Inst. Molekulargenetik, Univ.
 Mainz, Becherweg 32, D-55099 Mainz, GERMANY
 FEATURES
 source
 1. 234169
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="7"
 /clone="ICRFP703B1614Q5"
 /clone.lib="ICRFP703 + Research Genetics Mouse SAC-"
 join(20915..21109,22342..22464,28923..29048,34001..34135,
 39521..39646,42576..42692,43300..43389,48373..48489,
 49453..49575,53444..53560,54780..54902,55325..55441,
 57430..57534,61095..61250,61598..61804,75814..75984,
 77324..77491,77999..78160,80329..80526,81809..81961,
 85916..85995,87164..>87920)
 /gene="Cegpl"
 /evidence="experimental"
 20915..21109
 /gene="Cegpl"
 /number=1
 20915..87920
 /gene="Cegpl"
 join(20983..21109,22342..22464,28923..29048,34001..34135,
 39521..39646,42576..42692,43300..43389,48373..48489,
 49453..49575,53444..53560,54780..54902,55325..55441,
 57430..57534,61095..61250,61598..61804,75814..75984,

17324..77491,77999..78160,80329..80526,81809..81961,
 85916..85995,87164..87920)
 /gene="Cegpl"
 /codon_start=1
 /product="Cegpl protein"
 /protein_id="CA692293.1"
 /db_xref="GI:8052320"
 /db_xref="SPTREMBL:Q61J30"
 /translation="MSVAGGCRPREARAILLLILLPLLAAYVPDEGLINSPEDV
 DECAQDDCHADALCQNPISYKCSKPGTQGGPQEDXEDENILNGCCVHDCIN
 LFGNYRICIDGFMIADHGNCLDMDELENNGCCCHICINVISYECRKEGFFLSD
 NQHTCIHRSSEELSCNMKDHGCHICKEAPGSAACECPGFELAKWCKCILCHNG
 NGQOHCSEDIATGPECCHPRYLHADGSLGEGIVLEGTESNATSVADGKVK
 KRLIMETAVNNGGCDRTCKDTSTVHCSCPTGTLVDGKTCKDIDECUTANGCNH
 FCKNIVGSPCSCKKGFLLTDEKSCODVDECSLERICDHSCINHPGIFTCACNPYI
 LYSTHGGDTNEGSVNNGCCOQCNITVSGYECQCHPEKLNKKNKDCVKGPFELIS
 MIPVSLHCGSGCGDFLCRCRSGIHLSSDVTVIRISVFKNEGKCSLQKALISE
 GLPALPERHSSVKESFYANLITCSKQVPGALGRINAPMEFITVEFEIYKEV
 TASCNLSVAVRIEKLKRLKILKRAHREQFHLQSGMDLMAKIPSVSQHEET
 CGVGQHEESQVCSKAGTYDGSQRCILCPNGIFQNEESVTCFPCPKNELGSK
 ISEANVSDCGGLCQGEYSANGAPQOLCALGTFDVGRTSLSCGSLPIKHLGA
 ISFQDCETRVQCSGPHFYNTIHKICQPLGTYQPEFGKNVCSPGNTIIDFGGIN
 ITCKNKGSLGELGDFIGVIESPNYPGNYPANSECTMIINPPKRLILVVPFELPI
 FEDCGVLNKEKTSNNSVITYETQVSEPIAFISRSKLIQFKSNEGNSARGFQV
 PVIYDDYQELLIEDIVRGHLYASENHQELIKKALIKALFDVLAFHPQNYFYIAGE
 SREMPKPSFIRILLRSKVSRLPKK"
 21110..22341
 /gene="Cegpl"
 /number=1
 22342..22464
 /gene="Cegpl"
 /number=2
 22465..22622
 /gene="Cegpl"
 /number=2
 28923..29048
 /gene="Cegpl"
 /number=3
 29049..34000
 /gene="Cegpl"
 /number=3
 34001..34135
 /gene="Cegpl"
 /number=4
 34136..39520
 /gene="Cegpl"
 /number=4
 39521..39646
 /gene="Cegpl"
 /number=5
 39647..42575
 /gene="Cegpl"
 /number=5
 42576..42692
 /gene="Cegpl"
 /number=6
 42693..43299
 /gene="Cegpl"
 /number=6
 43300..43389
 /gene="Cegpl"
 /number=7
 43390..48372
 /gene="Cegpl"
 /number=7
 48373..48489
 /gene="Cegpl"
 /number=8
 48490..49452
 /gene="Cegpl"
 /number=8
 49453..49575
 /gene="Cegpl"

Wed Jul 3 08:56:42 2002

us-09-747-371-1.rge

Page 20

Job time: 10308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: July 2, 2002, 16:25:10 ; Search time 2733.88 seconds

(without alignments)

Sequence: 1 ggcgcgcgcgcacacccccc.....ctagacacttcggagacat 3737

Title: US-09-747-371-1

Perfect score: 3737

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

RST.*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estma:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_estci:**

9: gb_estli:**

10: gb_est2:**

11: gb_hic:**

12: gb_gss:**

13: em_gss_hum:**

14: em_gss_inv:**

15: em_gss_pin:**

16: em_gss_vrt:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715.4	45.9	2473	11	AK016571 Mus musc
2	672.6	18.0	809	10	BF526465
3	551	14.7	523	10	BM141778
4	550	14.7	530	10	BM142057
5	547.8	14.7	551	9	AW026325
6	544.8	14.6	869	10	BE542706
7	531.4	14.2	533	10	BE855427
8	528.4	14.1	530	10	BF196774
9	527.8	14.1	833	10	BE542707
10	520.4	13.9	573	10	BE871459
11	506.4	13.6	508	9	AA864639
12	497.8	13.3	501	10	BF330543
13	493.4	13.2	500	10	BF330548
14	483.6	12.9	493	9	AI287672
15	475	12.7	476	9	AI885416
16	470.6	12.6	478	9	AI424243
17	466.4	12.5	468	10	BE855854

C 18	452.1	12.4	648	9	AW577426
C 19	451	12.3	421	9	AI276664
C 20	450.6	12.4	552	10	BM030033
C 21	450	12.4	456	9	AI342436
C 22	450	12.4	457	9	AI420272
C 23	450	12.4	458	12	BE050881
C 24	450	12.4	459	10	BE858918
C 25	450	12.4	460	9	AI298488
C 26	450	12.4	461	9	AW602166
C 27	450	12.4	462	9	AI732365
C 28	450	12.4	463	10	BF589550
C 29	450	12.4	464	10	W79647
C 30	450	12.4	465	10	BF137295
C 31	450	12.4	466	9	AA689422
C 32	450	12.4	467	10	BE577567
C 33	450	12.4	468	9	AA228426
C 34	450	12.4	469	10	H05698
C 35	450	12.4	470	9	AI686213
C 36	450	12.4	471	9	AI420762
C 37	450	12.4	472	9	BE069582
C 38	450	12.4	473	9	AA224927
C 39	450	12.4	474	9	AI162869
C 40	450	12.4	475	9	AW339154
C 41	450	12.4	476	10	H93603
C 42	450	12.4	477	9	AA228353
C 43	450	12.4	478	9	AA574391
C 44	450	12.4	479	9	AA594242
C 45	450	12.4	480	9	BE661295

ALIGNMENTS

2473 bp mRNA linear HIC 19-JAN-2002

Mus musculus adult male testis cDNA, RIKEN full-length enriched

library, clone:493242019: similar to C63P1 PROTEIN, full insert

sequence.

ACCESSION AK016571

VERSION AK016571.1 GI:12855377

KEYWORDS CAP, trapper.

SOURCE Mus musculus (Strain: C57BL/6J) adult male testis cDNA to cDNA,

clone:493242019: full-length enriched mouse cDNA library,

clone:493242019.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (sites)

Cardinal P. and Hayashizaki Y.

High efficiency full-length cDNA cloning

Methods. Enzymol. 303, 19-44 (1999)

1 (sites)

Cardinal P., Shitaka Y., Hayatsu N., Sugawara Y., Shitaka K.

Identification and subtraction of cap-trapper selected clones to

prepare full-length cDNA libraries for rapid discovery of new genes

biochem. Res. 10 (10), 1617-1630 (2000)

1 (sites)

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

Shitaka K., Itoh M., Aizawa K., Nagata S., Sasaki M., Cardinal P.,

QY 1973 gggtatcgagaaacacaaatgtgtcagttgagggctgggaacattatgtatgagagacg 2032
 DB 1211 CGGTATGAGAAAGGACAGTGT---GTTGAGGCTGGGACGATATATGATGAGATCACA 1266
 QY 2033 agaacgtgtcattttatgtccaaatggaacacctccaaatgaggaaggaacaaatgacttt 2092
 DB 1267 AGAGCGCTGCAATTTGTGTCACAAAGGAAACCTTCACAAATGAGAGAGACAGACATG 1326
 QY 2093 tgaacattgcccagagacacgaattctggggccctgaagaccccgagagcttgaatat 2152
 DB 1327 TGACCATGCCCAAGACACAGAAATCTGGGTCTCTAAAATATCAGAAAGCTGGAATG 1386
 QY 2153 gtctaatgtgagaggtgtgtcaaacctgtgaatattcttcagatgggttgcaccttg 2212
 DB 1387 GTCTGACTGTGAGGACCTTTGTGAGCAGGTGAATTTTCAGAAAGATGATTTGCGCCCTG 1446
 QY 2213 ccagctctgtgcctgggacgtgtccagctgaagctgtgtggaacttctgttccctgt 2272
 DB 1447 CCAGCTGTGCTGGGACCTTTGAGCCTGTATGAGGCGGTACATCTTGCCCTCAIG 1506
 QY 2273 tggagagagccttgcacacaaacatcagggagctacttcttcagagctgtgaacacag 2332
 DB 1507 CGGAGGAGGTCTTCCACCAAAATCTGGGACCCACTTCTTCCAGGATTTGAAACCG 1566
 QY 2333 agtcaattgtccactggaactttctaacacacacacactcaccgagtattctgttgc 2392
 DB 1567 AGTCAATGCTCGCTGGGACCTTTACACACACACACACACAGATGATCCATGCGCC 1626
 QY 2393 agtgggaacataccagcctgatttggaaaaataattgtttcttgcacaggaatatc 2452
 DB 1627 ACTAGGAGCATACAGCCTGATTTGAAAAATAATTTGTTTCTGCGCGGCAATAC 1686
 QY 2453 taacactgcttgatgtctccaaacataaaccaagtgtaaaacagagatgtggag 2512
 DB 1687 TACCACGAGCTTGTGCTGCTCCACACATAACACAGTGTAAACACAGATGAGTGGAG 1746
 QY 2513 ggaactggagatttcaactgggtacattgattcccaaaactaccagcgaattaccagc 2572
 DB 1747 GGACTGGGGATTTACAGGGGTATTTGATCCCAAACTACCTTGGGAATACCCAGC 1806
 QY 2573 caacacagagtgtaogtggagacataacacacacacacacacacacacacacacacac 2632
 DB 1807 CAACTCAGAGTGCACCTGACCATCAACCCACCCACCCACCCACCCACCCACCCACCC 1866
 QY 2633 cctggaattctcctccatagagagcactgtgggactatctgtgtgagggagaaac 2692
 DB 1867 CCGTGAATCTTCTGCCATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1926
 QY 2693 ctctccaaattctgtgacacatatgaaacctgccagactacgaaccccatcgc 2752
 DB 1927 CTTCTCATCAATCTGTGACACATATGAAACCTGCCAGCTTATGAAAGCCCACTCCG 1986
 QY 2753 ctccacactccaggtcaaaagactgtgattcagttcagttcagttcagttcagttcagtt 2812
 DB 1987 CTTCACTCCAGGTCAAAAGAACTGTGGATTTCAGTTCAAGTCAATGAAGGAAACAGTGC 2046
 QY 2813 tagaggttccaggtcccatcactgaactatgatgatgatgatgatgatgatgatgatga 2872
 DB 2047 CCGGGGGTTCCAGGTCTCCCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2106
 QY 2873 catagtccagatggcaggtctctatgctctgagaacacacacacacacacacacacacac 2932
 DB 2107 CATAGTTCGATGCGAGCTCTATGCTCTGAGAACCAATCAGAAATATCAGAGATAA 2166
 QY 2933 gaaacttcaaggtctgttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtct 2992
 DB 2167 GAAACTGTATCAAGCTCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226
 QY 2993 agccagaggtccgagagatgttccagatctgtccagatctgtccagatctgtccagat 3052
 DB 2227 AGCCAGGAGTCCCGAGAGATGTTTCCAGATCTTTCCAGATCTTTCCAGATCTTTCCAG 2286
 QY 3053 gtccaggtttttgagacttacaatgactcagccacgtgcccactcactcaatacaaatgtt 3112

DB 2287 GTCAAAATTTTTCAGGAGGATTAACAAATGACATCTTCCACACGCGATGTGGAC-AGGGTCC 2345
 QY 3113 tctctcagaggtgtgtgtgagagaggtgtctctctctctctctctctctctctctctctct 3172
 DB 2346 TACTTGTGAGTTCAGTGGGCTGACAGCGGTTTCCCTTCACATG-CCGCTTCCATGCGAT 2404
 QY 3173 gctgtctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 3232
 DB 2405 ATTACACACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 2464
 QY 3233 gctaaactg 3241
 DB 2465 GATAAATG 2473

RESULT 2
 BF526465 809 bp mRNA linear EST 11-DEC-2000
 LOCUS 60207057F1 NC1_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4213550
 DEFINITION 5' mRNA sequence.
 ACCESSION BF526465
 VERSION BF526465.1 GI:11613828
 KEYWORDS EST.
 ORGANISM Homo sapiens
 Human.
 REFERENCE 1 (cases 1 to 509)
 AUTHORS NIH-N32 http://hgc.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contract: Robert Strausberg, Ph.D.
 Email: g3apb-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L14M9786 row: b column: 15
 High quality sequence stop: 746.
 Location/Qualifiers
 source 1..809
 organism="Homo sapiens"
 db_xref="taxon:9606"
 clone="IMAGE:4213550"
 /clone_lib="NC1_CGAP_Brn64"
 /tissue_type="glioblastoma with EGFR amplification"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: Brain; Vector: PCMV-SPORE6; Site: 1: NC1; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NC1_CGAP library."

BASE COUNT 227 a 204 c 193 g 185 t
 ORIGIN

Query Match 18.0%; Score 672.6; DB 10; Length 809;
 Best Local Similarity 95.4%; Pred. No. 6.3e-115;
 Matches 747; Conservative 0; Mismatches 29; Indels 7; Gaps 5;

QY 1968 ggccagggctcagcagaaacacaaatgtctcagttgcggggctgggacattatgtatga 2027
 DB 3 GGGCAAGTCAAGCAGAAACCAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 62
 QY 2028 gacacgagagctcattttatgtccaaatggaaccttccaaatgaggaagacaaatg 2087
 DB 63 GCAAGAGAGCTGATTTTATGTCAAATGGAGCTTCCCAAAAGAGAGAGAGAGAGAG 122
 QY 2088 actatgaaacatgcccaagacacaggaattcttgggacctgagagaccccaagacttgg 2147

123 ACTGTGACACATCCCAAGACAGAAATTCCTGGGCGCTGAGACCCCAAGACCTTGG 182
 QY 2148 aatattctgtaattgagtgctgtctcaacctggtgaattattctcagatggtctgca 2207
 DB 183 AATATGCTGAATGGAGGTCTGTCAACCTGTGAATATTCGACAGTGGCTTGGCA 242
 QY 2208 cctgcagctctgtgacctggcagcttccagctcaaacctgagctgagcttctctgttc 2267
 DB 243 CCTTGCCACTGTGCTGGCAGCTTCCAGCTGAGCTGTGAGACTTCTGCTTTC 302
 QY 2268 cctgtggaggagccttgcacacacacacacacacacacacacacacacacacac 2327
 DB 303 CCTGTGGAGGAGGCTTGGCACAACATCAGGAGCTATCTCTTTCAGACATGAA 362
 QY 2328 accagagttcaagtccaactgacattctcaacacacacacacacacacacacacac 2387
 DB 363 ACCAGATTCAATGTCACTGGACATTTCTACACACACACACACACACACACACAC 422
 QY 2388 tgcctggtgggaacataccagcctgaatttggaaaaataattgtttcttctgcccagga 2447
 DB 423 TGCCAGTGGGAGACATACCACTGAAATTTGGAAAAAATAA-TGTGTCTTCTGGCCAGCA 481
 QY 2448 aatacacacagcagcttggctgcacacacacacacacacacacacacacacacacacac 2507
 DB 482 AATACACAGCAGCTTGTGCTGCTCCACACACATACCCAGTGTAAACACAGAAATGT 541
 QY 2508 ggaaggagctgggagatttcaactggtgacattgaatcccc-aaactacccaggaatt- 2565
 DB 542 GGAGGGAGCTGGGAGATTTCACTGGGTACATTGAATCCCAAACTACCCAGGCAAAAT 601
 QY 2566 -accagccacacacagctgtagcggacacacacacacacacacacacacacacacacacac 2623
 DB 602 ACCCAGCAACACACAGCTGAGTGGACCATACACACACACACACACACACACACAC 660
 QY 2624 gatcgtgctccctgagatttctctgcccataagaggagcagctgtgggagctattctggtgat 2683
 DB 661 GATCGTGTCTCTGAGATCTTCTGTCATAGAGGAGCAGCTGTGGGAGCTATCTGGTCAI 720
 QY 2684 gcggaaaacctcttcacattctgtgacacacacacacacacacacacacacacacacacac 2743
 DB 721 CGGACACACTCTTCTCAATCTGTGACAC-TATGAACCTGTGACAAATATGGAAG 779
 QY 2744 ccc 2746
 DB 780 CCC 782
 RESULT 3
 BML1778/c
 LOCUS 1125b09.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 DEFINITION cDNA 3', mRNA sequence.
 ACCESSION BML1778
 VERSION BML1778.1 GI:17151843
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 553)
 AUTHORS Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kastner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bristain, R.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, W., Williams, J.,
 Jackson, Y., and Bowers, Y.
 TITLE Endocrine Pancreas Consortium
 JOURNAL Unpublished (2000)
 COMMENT Contact: Douglas Melton, Klaus H. Kastner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138

123 51-195-1812
 FAX: 617-495-8583
 Email: dmelton@comp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center for information on
 ordering a clone please contact: Juliana Brown
 (jbrown@wustl.edu)
 Seq primer: -400p from 3'ccg
 RFL quality sequence stop: 441.
 Location/Qualifiers
 1..553
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /note="Jordan: Pancreas: Vector: pSPORT1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library Kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column.
 Fractionation: average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product.
 representing library inserts and hybridized to an EcoT of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."
 BASE COUNT 145 a 118 c 133 g 157 t
 CRIGIN
 Query Match 14.7%; Score 551; DB 10; Length 553;
 Best Local Similarity 100.0%; Pred No. 2.4e-92;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2693 cctcttctatcattctgtgacacacatgaacctgcagagcctacacagcccccacgc 2752
 DB 553 CTCCTATCCGATTTCTGACAAATATGAAGCTGCCAGACCTACGAGCCCTCCATCCG 494
 QY 2753 ctctcctctcagctcaagagctgtgattcagctcagctccacacacacacacacacac 2812
 DB 493 CTTCATCTCCAGTCAAGACAGCTGTGATTCATTCATGATCCATCAATGAGGACACAC 434
 QY 2813 tagaggtttccaggtcccatcagtgacatatgatgaggactaccaggtgactcattgaga 2872
 DB 433 TAGAGTGTTCAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATG 374
 QY 2873 catatctcagatggtcagctctatgagctgagacacacacacacacacacacacac 2932
 DB 373 CATATCTCAGTCAAGACAGCTGTGATTCATTCATGATCCATCAATGAGGACACAC 314
 QY 2933 gaaacacacacaggtctcttctgagctcctgggcccacacacacacacacacacacac 2992
 DB 313 GAAACATCTCAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATG 254
 QY 2993 agcctcagctcagcagac 3052
 DB 253 AGCCTCAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 194
 QY 3053 gtctcaggtttctgagcctcacaagtactcagcagcagctgcacacacacacacacac 3112
 DB 193 GTCTCAGTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 134
 QY 3113 tgcctacacagctgtgtggacacagctgtctctctctctctctctctctctctctctct 3172
 DB 133 TGCCTACAGCTGTGAGGACAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 74
 QY 3173 gctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 3232

ORGANISM	HOMO SAPIENS
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; HOMO.
AUTHORS	1 (bases 1 to 573) Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Bfiones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RG3&t2=SC3-FR0141-060700-021-n05&t3=2000-07-06&t4=1) Seq primer: puc 18 forward High quality sequence start: 15 High quality sequence stop: 573. Location/Qualifiers 1..573 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="FR0143" /dev_stage="Adult" /note="Organ: prostate_normal; Vector: puc18; Site:1: Small ; Site:2: Small; A mini-library was made by cloning products derived from ORESNES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
FEATURES	143 a 128 c 167 g 135 t
SOURCE	143 a 128 c 167 g 135 t
BASE COUNT	
ORIGIN	

Q7	1274	caaccactctgagagacaccccaatgagtgcaagcatcaacaacgagggagctgtccacaggtctg	1333
Q7	1275		
Q7	1276		
Q7	1277		
Q7	1278		
Q7	1279		
Q7	1280		
Q7	1281		
Q7	1282		
Q7	1283		
Q7	1284		
Q7	1285		
Q7	1286		
Q7	1287		
Q7	1288		
Q7	1289		
Q7	1290		
Q7	1291		
Q7	1292		
Q7	1293		
Q7	1294		
Q7	1295		
Q7	1296		
Q7	1297		
Q7	1298		
Q7	1299		
Q7	1300		
Q7	1301		
Q7	1302		
Q7	1303		
Q7	1304		
Q7	1305		
Q7	1306		
Q7	1307		
Q7	1308		
Q7	1309		
Q7	1310		
Q7	1311		
Q7	1312		
Q7	1313		
Q7	1314		
Q7	1315		
Q7	1316		
Q7	1317		
Q7	1318		
Q7	1319		
Q7	1320		
Q7	1321		
Q7	1322		
Q7	1323		
Q7	1324		
Q7	1325		
Q7	1326		
Q7	1327		
Q7	1328		
Q7	1329		
Q7	1330		
Q7	1331		
Q7	1332		
Q7	1333		
Q7	1334		
Q7	1335		
Q7	1336		
Q7	1337		
Q7	1338		
Q7	1339		
Q7	1340		
Q7	1341		
Q7	1342		
Q7	1343		
Q7	1344		
Q7	1345		
Q7	1346		
Q7	1347		
Q7	1348		
Q7	1349		
Q7	1350		
Q7	1351		
Q7	1352		
Q7	1353		
Q7	1354		
Q7	1355		
Q7	1356		
Q7	1357		
Q7	1358		
Q7	1359		
Q7	1360		
Q7	1361		
Q7	1362		
Q7	1363		
Q7	1364		
Q7	1365		
Q7	1366		
Q7	1367		
Q7	1368		
Q7	1369		
Q7	1370		
Q7	1371		
Q7	1372		
Q7	1373		
Q7	1374		
Q7	1375		
Q7	1376		
Q7	1377		
Q7	1378		
Q7	1379		
Q7	1380		
Q7	1381		
Q7	1382		
Q7	1383		
Q7	1384		
Q7	1385		
Q7	1386		
Q7	1387		
Q7	1388		
Q7	1389		
Q7	1390		
Q7	1391		
Q7	1392		
Q7	1393		
Q7	1394		
Q7	1395		
Q7	1396		</

sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ctl=RC4f2-RC4-ET0337-
 180500-015-h048f3-2000-05-18&ct=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 500.
 Location/Qualifiers
 1..500
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT0337"
 /dev_stage="Adult"
 /note="Organ: breast; Vector: puc18; Site:1: Smar; Site:2:
 Smar; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 142 a 127 c 131 g 100 t
 ORIGIN

Query Match 13.28; Score 493.4; DB 10; Length 500;
 Best Local Similarity 99.8%; Pred. No. 1.2e-81;
 Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1554 accttaagctaaatgaagcaagctgttagtttgaataatgtcgtgctgtttcccgagggc 1613
 DB 6 ACCTTTAAGCTAAATGAGGCAAGTGTAGTTTGAATAATGCTGAGCTGTTCGCGAGGCT 65
 QY 1614 ctgcgaccgactaccgagagacacagctcagtaaaagagagcttccgctactyaaac 1673
 DB 60 CTGCGACCACTACCAGAGACACAGCTCAGTAAAGAGAGGCTTCGCTAGTAAAC 125
 QY 1674 cttacatgcagctctgcgaagcaagctccagagagccctggtgcgacccaagccctaaag 1733
 DB 136 CTTACATGCAAGCTCTGGCAAGCAAGTCCCGAGGAGCCCTGGCGACCAAGCACCCCTAAG 185
 QY 1734 gaaatgtttatcacctgttgagtttgagcttgaaactaaacaaagagagtgacagcttct 1793
 DB 186 GAAATGTTTATCAGTGTGAGTTGAGCTTGAACCTAACCAAGAGAGTGACAGCTTCT 245
 QY 1794 tgtgacctgagctgcctaatgaagacacagagagcgtctcctytaaagccatccgcagc 1853
 DB 246 TGTGACCTGAGCTGCTATGAAAGCAAGCAAGCGGCTCCGTAAGACCCCAATG 305
 QY 1854 ctcaagaagccctccagagagcagtttcaacctccagctctcaggtcgaatgaacctgcag 1913
 DB 306 CTCAGAAGGCCGCTCCAGAGGAGCAGTTTCACCTCCAGCTCTCAGGATGAGCCGAC 365
 QY 1914 gtggctaaaagcctcccgaaacatctgaacgcagcagcagctcctgtggagtgggcccag 1973
 DB 366 GTGGCTAAAGAGCTCCGAGACATCTGAGCGCCAGCAGAGTCTGTGGAGTGGGCCAG 425
 QY 1974 ggtcatgcagaaaacaaatgtctcagttgaggggtggacacctattatgatggagcagca 2033
 DB 426 GGTCAATGCAGAAAACAAATGTCTCAGTTGTCAGTGGAGGCTGGGACCTATTATGATGGAGCAGA 485
 QY 2034 gaacgcgtcatttta 2048
 DB 486 GAACGCTGCATTTTA 500

RESULT 14
 A1287672/c
 LOCUS
 DEFINITION
 3 mRNA sequence.
 A1287672
 A1287672.1 GI:19271425
 EST.
 Human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 493)
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cstra@rmail.nih.gov
 This clone is available royalty-free through LHM. Contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: 430P from G100
 High quality sequence stop: 418.
 Location/Qualifiers
 1..493
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Scars_MHMPv1_S1"
 /clone_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10b"
 /note="Organ: mixed (see below); Vector: pTZ19-Pac
 (Pharmacia) with a modified polylinker; Site:1: Not i;
 Site:2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2M9M, pregnant uterus
 MHMP, and fetal heart NDH19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 26032-26523,
 34048-345179, and 484188-489479."
 BASE COUNT 119 a 125 c 120 g 124 t 1 others
 ORIGIN
 Query Match 12.9%; Score 483.6; DB 9; Length 493;
 Best Local Similarity 99.0%; Pred. No. 7.9e-80;
 Matches 485; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 3246 tggctttttcccgagacatcgtgagatgagatgagatgagatgagatgagatgagatgagat 3305
 DB 492 TGGCTTTTTCCTCCCGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 453
 QY 3306 tctcactgtctgtggtcgtgtctgtgagatgacacggtggtggtggtggtggtggtggt 3365
 DB 432 TCTTACTGCTGTGTGGCGGTGTCTTGGTAAGTACAGGGCTGGCTTACAGGAGTCTTGGT 373
 QY 3366 cagcttgggtggaactcactgtctctgtggtgttactctctcactcaggtggtctgag 3425
 DB 372 CAGGCTTGGTGGAGTACCTGCTGCTGCGGGTCTTACCTTCTCAGGAGTCTTGGT 313
 QY 3426 tggaaagagggcacagataaagctgtattctgaaactcagcttctctagccggc 3485
 DB 312 TGGAAAGAGGGCCACAGATAAAGCTGCTTATTCTGAACCTTACGCTTCTTACCCCGCC 253
 QY 3486 cctctctaaagggagccctctgcactcgtgtgcaggtctgtgacccaggtcagacaggaaga 3545
 DB 252 CCTCTCTAAGGGAGGCCCTCTGCACCTCGTGTGTCAGGCTCTGACAGGCTCTGACAGGCTCAGG 193

Wed Jul 3 08:56:44 2002

us-09-747-371-1.rst

Page 13

QY 2530 ctgggtacattgaatcccaactaccaggaattaccacagccaaca 2577
|||||
DB 194 ctgggtacattgaatcccaactaccaggaattaccacagccaaca 241

RESULT 2
US-08-897-443-2
Sequence 2, Application US/08897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0348 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3373 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UTRSN02
CLONE: 681719
US-08-897-443-2

Query Match 2.2%; Score 83; DB 2; Length 3373;
Best Local Similarity 50.7%; Pred. No. 9e-12;
Matches 253; Conservative 0; Mismatches 240; Indels 6; Gaps 2;
QY 925 tcatggaacgtgtgcttcaacaatgagggtgtgacccgacccgtgaaggaatcttga 984
|||
DB 890 TCAGAGATCTGTGTGCCATGCGAGACACCACTGTGACGACCTCTGTGTGAATGTCCGG 949
QY 985 caggtgtccactgcagttgtccctgttggtattccattccagttggtggaagacatga 1044
|||
DB 950 GCTCTTGTCTGTCAGTGTACAGTGGCTAGCTGGCTAGCCCTGGCTGAGGATGGGAAGAGTGTG 1009
QY 1045 agatattgatgtgacagacacccgcaatgaggtgtgtgatcattcttcgaaaacatcg 1104
|||
DB 1010 TGCTGTGGACTACTGTGTGCTCTCAAAACACCGGATGTGCACATGAGTGTGTAATGTGTG 1059
QY 1105 tggcaggttttgactggtggtgtgcaagaaggatttaataattacacagatggaagtctt 1164
|||
DB 1070 ATGCTCTTACCTTTGGCAGTGGCCATCAAGGATTGTCTTACCCAGATATAAAACGT 1129

QY 1165 gcaaacgtgtggtatgagtg---ctcttggataggacctgtgaccacagctgcatcaacc 1221
|||
DB 1130 GCAAAAGATAGACTACTGTCCCTCACTAATACAGGATGTCACAGAGTGTGTGAACA 1189
QY 1222 accctgtgacattgtctgtgcaaccgaggtgtacacctgtat---ggcttcaccc 1278
|||
DB 1150 CAGATGATTCCTATTCCISCCACTGCCIGAAGGCTTTACCTGAATCCAGATAAGAAA 1249
QY 1279 actgtgtgagacacccaatgtgtgagcatcaacaacgaggtgtgacagctgtgtgtga 1338
|||
DB 1250 CCTGTGAGAGATCAACTACTGTGCACTGACAAACCGGCTGTGTGAGATGATGGGTGA 1309
QY 1339 acacagtggtgagctatgaagccacccctggtgtacaaagctccactggaataaaa 1398
|||
DB 1310 ACAATGGAGAGACTACTAGCCGCTGCCACCGTGGCTTACACTGTGACACCCCAATGGCA 1369
QY 1399 aagactgtgtggaagtga 1417
|||
DB 1370 AAATCTGACGCGGAGTGA 1388

RESULT 3
US-08-833-963C-1
Sequence 1, Application US/08833963C
Patent No. 5916769
GENERAL INFORMATION:
APPLICANT: Olsen, et al.
TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
TITLE OF INVENTION: H2AB58X
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Ave
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.36
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,963C
FILING DATE: 11-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/05033
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brooks, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 125..1453
US-08-833-963C-1

Query Match 2.1%; Score 78.6; DB 2; Length 1531;
Best Local Similarity 54.6%; Pred. No. 8.4e-11;
Matches 206; Conservative 0; Mismatches 159; Indels 12; Gaps 2;

```

RESULTS
US-09-188-930-67
: Sequence 67, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
: APPLICANT: Murison, James Grey
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FastSeq For Windows Version 3.0
: SEQ ID NO 67
: LENGTH: 1250
: TYPE: DNA
: ORGANISM: Rat
US-09-188-930-67

Query Match      2.0%; Score 75; DB 3; Length 1250;
Best Local Similarity 51.8%; Prd. No. 6.6e-10;
Matches 184; Conservative 4; Mismatches 164; Indels 3; Gaps

Qy 1068 cgcacagagaggttggatcattttgcgaacacatcgtggcgagctttttactgcgcgtgc 118
      ||| ||| || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 cgcataggctacgcgcagcgtctgtgcgaatgttcgtgattccattccctgcagctgt 215

Qy 1128 agaaagagagattcaattattacagatagaaagctttgccaaagatgttgatgagctct 118
      || | ||| || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 aaccttgcttcacctcaacgagatgagagagctttgtgccaaagatgagacagagtgaa 305

```



```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Wed Jul 3 08:56:43 2002

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267,387
FILING DATE: 29-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93 201 906.0
FILING DATE: 30-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B038633
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3290 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-267-387-1

Query Match 1.6%; Score 58.8; DB 1; Length 3290;
Best Local Similarity 54.3%; Pred. No. 1.8e-05;
Matches 164; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 948 aatgaggtgtgacgcacgtgtaagagatacttcgacaggtgtccactgcagttgtcct 1007
DB 621 AATGAGGTTGTCAGTCAATTTGTGATATACACCTGGAGTACCACCTGTTCTGTAAA 680
QY 1008 gttggtattcactctccagttggtgagagacatgtaagagatattgagagtgcca---g 1064
DB 681 AATGTTTGTGTTGCTTCAATTAAGAAAGATTGTAAGAGATGGGATCAATGCTCTTTG 740
QY 1065 accgcgaatgaggtgtgtgatcattctgcgaaacacatctgtggcagttttgactgcgc 1124
DB 741 AAGCCAGACATTTGTGTCACACTGTGTCAGAGACATCCAGGAGATTTTGAATGTGAA 800
QY 1125 tqcaagaaagatttaattacacagattgagaggtcttggccaagattggtgagtgagtc 1184
DB 801 TCCCCGGAAGGCTACAGATATATCTCAATCAAAAGTCTGTGAGATATAGATGAATGC 860
QY 1185 tcttgataggacctgtgacacagctgcatcaacacacccctggcacatttgcctgtgct 1244
DB 861 TCT---GAGAACATGTGTCTCAGCTTTGTCTCAATTACCTTGGAGGTACACTTGTCTAI 917
QY 1245 tg 1246
DB 918 TG 919

RESULT 15
US-08-651-472-67/c
Sequence 67, Application US/08651472
Patent No. 6103244
GENERAL INFORMATION:
APPLICANT: DORNER, Friedrich
APPLICANT: SCHEIFFLINGER, Friedrich
APPLICANT: FALKNER, Falko Gunter
APPLICANT: PFELEIDERER, Michael
TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
TITLE OF INVENTION: (HIV-1) ANTIGENS
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,472
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,738
FILING DATE: 20-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,080
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENI, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/166/IMK0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 6811 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pM2-gpta Prots
US-08-651-472-67

Query Match 1.6%; Score 58.8; DB 3; Length 6811;
Best Local Similarity 54.3%; Pred. No. 2.6e-05;
Matches 164; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 948 aatgaggtgtgacgcacgtgtaagagatacttcgacaggtgtccactgcagttgtcct 1007
DB 4437 AATGAGGTTGTCAGTCAATTTGTGATATACACCTGGAGTACCACCTGTTCTGTAAA 4378
QY 1008 gttggtattcactctccagttggtgagagacatgtaagagatattgagagtgcca---g 1064
DB 4377 AATGTTTGTGTTGCTTCAATTAAGAAAGATTGTAAGAGATGGGATCAATGCTCTTTG 4318
QY 1065 accgcgaatgaggtgtgtgatcattctgcgaaacacatctgtggcagttttgactgcgc 1124
DB 4317 AAGCCAGACATTTGTGTCACACTGTGTCAGAGACATCCAGGAGATTTTGAATGTGAA 4256
QY 1125 tqcaagaaagatttaattacacagattgagaggtcttggccaagattggtgagtgagtc 1184
DB 4257 TCCCCGGAAGGCTACAGATATATCTCAATCAAAAGTCTGTGAGATATAGATGAATGC 4198
QY 1185 tcttgataggacctgtgacacagctgcatcaacacacccctggcacatttgcctgtgct 1244
DB 4197 TCT---GAGAACATGTGTCTCAGCTTTGTCTCAATTACCTTGGAGGTACACTTGTCTAI 4141
QY 1245 tg 1246
DB 4140 TG 4139

Search completed: July 2, 2002, 19:19:37
Job time: 7613 sec

